Bounds and Error Estimates for Radiosity

Dani Lischinski *Brian Smits* *Donald P. Greenberg*

Program of Computer Graphics*
Cornell University

ABSTRACT

We present a method for determining *a posteriori* bounds and estimates for local and total errors in radiosity solutions. The ability to obtain bounds and estimates for the total error is crucial for reliably judging the acceptability of a solution. Realistic estimates of the local error improve the efficiency of adaptive radiosity algorithms, such as hierarchical radiosity, by indicating where adaptive refinement is necessary.

First, we describe a hierarchical radiosity algorithm that computes conservative lower and upper bounds on the exact radiosity function, as well as on the approximate solution. These bounds account for the propagation of errors due to interreflections, and provide a conservative upper bound on the error. We also describe a non-conservative version of the same algorithm that is capable of computing tighter bounds, from which more realistic error estimates can be obtained. Finally, we derive an expression for the effect of a particular interaction on the total error. This yields a new error-driven refinement strategy for hierarchical radiosity, which is shown to be superior to brightness-weighted refinement.

CR Categories and Subject Descriptors: I.3.3—[**Computer Graphics**]: Picture/Image Generation; I.3.7—[**Computer Graphics**]: Three-Dimensional Graphics and Realism – Radiosity.

Additional Key Words and Phrases: *a posteriori* error bounds and estimates, adaptive refinement, global illumination, hierarchical radiosity, importance

1 INTRODUCTION

During the past decade, radiosity has become the method of choice for simulating global illumination in environments consisting entirely of Lambertian (ideal diffuse) reflectors and emitters. Global illumination in such environments is governed by a Fredholm integral equation of the second kind:

$$B(x) = E(x) + \rho(x) \int k(x, y)B(y)\, dy, \qquad (1)$$

where $B(x)$ is the total radiosity at point x, $E(x)$ is the emission, and $\rho(x)$ is the reflectivity[1]. The kernel of the integral operator

$$k(x, y) = \frac{\cos\theta_x \cos\theta_y}{\pi r_{xy}^2} V(x, y)$$

describes the point-to-point transfer from y to x: θ_x and θ_y are the angles between the surface normals and the line connecting x and y, r_{xy} is the distance between the two points, and $V(x, y)$ is 1 if y is visible to x and 0 otherwise.

Radiosity algorithms compute an approximate solution $\widehat{B}(x)$ to equation (1) by projecting the continuous integral equation into a finite dimensional function space. Most radiosity algorithms use spaces of piecewise constant functions, although recently several investigators described radiosity algorithms that use higher order basis functions [12, 21, 22]. The projection of the continuous equation results in a discrete system of n linear equations, where n is the dimension of the finite dimensional space:

$$(I - K)B = E. \qquad (2)$$

Since radiosity algorithms produce only approximate results, reliable error bounds and realistic error estimates are crucial for assessing the acceptability of a particular solution, as well as for automatic adaptive refinement.

The theory of integral equations [1, 10] provides an *a priori* error analysis of computational methods for Fredholm equations of the second kind. This general framework is valuable for analyzing convergence rates [10], and for bounding various components of the error in terms of the corresponding operator norms [3]. However, *a priori* error bounds are typically pessimistic and often difficult to evaluate, requiring information about the operators or the exact solution, which may not be available. In this paper we discuss *a posteriori* methods that bound the combined effect of the various error sources for a particular instance of the problem. Because the analysis takes place after or during the computation of the solution, such methods can provide tighter bounds more easily.

During the 1980's the issues of *a posteriori* error estimation and automatic adaptive refinement received considerable attention in the finite elements literature [4, 19]. Most of the work has related to partial differential equations. More recently, results have become available also for boundary integral equations and boundary element methods [5], which are essentially similar to radiosity methods.

In the radiosity literature, however, very little has been written regarding *a posteriori* error analysis. In particular, current radiosity algorithms can provide the user with neither guaranteed bounds on the total error $\|B(x) - \widehat{B}(x)\|$, for any function norm $\|\cdot\|$, nor reliable

* 580 ETC Building, Cornell University, Ithaca, New York 14853. E-mail: {danix | bes | dpg}@graphics.cornell.edu

[1] All of these quantities depend on the wavelength, but for simplicity we restrict the discussion to a single wavelength throughout this paper.

estimates of the total error. The goal of this paper is to address these deficiencies.

We begin with a brief account of past attempts to estimate the error in radiosity approximations.

In Section 3 we describe an algorithm that computes two piecewise constant functions that are guaranteed to bound the exact solution $B(x)$ from below and from above. This can be done simultaneously with the computation of the approximate solution $\widehat{B}(x)$. These bounds pertain to both $B(x)$ and $\widehat{B}(x)$, so they can be used to compute a conservative upper bound on the error $\|B(x) - \widehat{B}(x)\|$. Our algorithm resembles the general bracketing technique suggested by Brown [6], but we utilize special properties of equation (1) to produce tighter bounds.

In many instances, a conservative error bound is not required, or is too pessimistic, and a good error estimate is preferable. In Section 4 we describe a non-conservative bounding algorithm that computes tighter "bounds", which are no longer guaranteed to contain the exact solution $B(x)$ everywhere. These non-conservative bounds are much cheaper to compute, and result in more realistic error estimates, i.e., estimates which are closer to the actual errors.

Our final contribution is the derivation of an expression that describes the effect that a particular interaction has on the total error bound. This expression uses bounds on importance and on radiosity that can be computed using our algorithms. This gives rise to a new error-driven refinement strategy for hierarchical radiosity, which is shown to be superior to brightness-weighted refinement.

2 ERROR ESTIMATION IN RADIOSITY

Cohen's adaptive subdivision algorithm [8] can be viewed as the first attempt at *a posteriori* error estimation. In this algorithm, elements are subdivided if their radiosity differs from that of their neighbors by more than a certain threshold. Thus, the difference between radiosities of adjacent elements is treated as an indicator of the local error. Note that this heuristic may fail to identify elements that should be subdivided.

A more conservative approach is described by Campbell [7], who uses numerical optimization to find the minimum and the maximum radiosity over each element. The difference between these extrema is used as a criterion for adaptive subdivision.

Among the various radiosity methods, hierarchical radiosity (HR) comes closest to bounding the errors in the solution: a bound is computed on the error in each interaction (link) between patches. Hanrahan et al. [13] compute an approximate upper bound on the form-factor corresponding to the link, and use it to obtain an approximate upper bound on the transferred energy.

Smits et al. [17] obtain an estimate of the error in each interaction by point sampling the kernel function over the areas of the two interacting patches. The difference between the maximum and the minimum values thus obtained is taken to be the error in the form-factor. This error estimate is less conservative than Hanrahan's, but it is generally more accurate.

Given error estimates for all the links in an environment, one can approximate the total error over each patch by summing up the errors associated with all its links. However, this does not take into account propagation of errors, which parallels the propagation of light.

Several other error estimation techniques are surveyed by Cohen and Wallace [9]. These techniques can be classified into two groups: (i) comparing the approximation to a higher order approximation, and (ii) computing the residual

$$r(x) = \widehat{B}(x) - E(x) - \rho(x) \int k(x,y)\widehat{B}(y)\, dy.$$

Both of the above are computationally expensive and do not provide foolproof subdivision criteria.

3 A RADIOSITY-BOUNDING ALGORITHM

For simplicity we start with a radiosity-bounding algorithm for full matrix radiosity. Later in this section, we extend this algorithm to work within a more efficient radiosity algorithm, namely hierarchical radiosity.

The goal of the bounding algorithm is to compute two piecewise constant functions $\underline{B}(x)$ and $\overline{B}(x)$ that bound the exact radiosity $B(x)$ from below and from above, respectively. In other words, given a discretization of the environment into a set of n elements $S_1, \ldots, S_n$, we compute $\underline{B}_1, \ldots, \underline{B}_n$ and $\overline{B}_1, \ldots, \overline{B}_n$, such that

$$\underline{B}_i \leq B(x) \leq \overline{B}_i \quad \text{for all } x \in S_i.$$

Roughly, this is achieved by replacing the coefficients of the discrete linear system (2) by infima and suprema bounding the corresponding continuous functions over the areas of the elements $S_1, \ldots, S_n$.

3.1 Computing Lower Bounds

Let $\underline{\rho}_i$ and $\underline{E}_i$ denote the infima on the reflectivity and on the emission, respectively, over S_i. Also, let $\underline{F}_{ij}$ denote the infimum on the (point-to-area) form-factor from a point on S_i to S_j:

$$\underline{F}_{ij} = \inf_{x \in S_i} \int_{S_j} k(x,y)\, dy,$$

and let $\underline{K}_{ij} = \underline{\rho}_i \underline{F}_{ij}$. Then the lower bounds vector $\underline{B} = (\underline{B}_1, \ldots, \underline{B}_n)^{\mathrm{T}}$ can be obtained by solving a linear system of n equations

$$(I - \underline{K})\, \underline{B} = \underline{E},$$

using a standard method such as Jacobi or Gauss-Seidel iteration.

Proof. It is easy to see that the matrix $I - \underline{K}$ is strictly diagonally dominant, from which it follows [11] that the spectral radius of $\underline{K}$ is strictly less than 1. Thus, the inverse matrix $(I - \underline{K})^{-1}$ exists, and can be expressed as a Neumann series

$$(I - \underline{K})^{-1} = I + \sum_{i=1}^{\infty} \underline{K}^i.$$

The solution vector $\underline{B}$ can then be written as

$$\underline{B} = (I - \underline{K})^{-1}\underline{E} = \underline{E} + \sum_{i=1}^{\infty} \underline{K}^i \underline{E}.$$

Alternatively, $\underline{B}$ can be expressed as the limit of the series of vectors $\underline{B}^{(k)}$ that are defined by the recurrence relation

$$\begin{aligned} \underline{B}^{(0)} &= \underline{E} \\ \underline{B}^{(k+1)} &= \underline{E} + \underline{K}\underline{B}^{(k)}. \end{aligned}$$

Note that the above corresponds to Jacobi iteration in matrix notation. It is easy to see by induction that $\underline{B}^{(k)}$ is a lower bound on the radiosity after k light bounces. Therefore, the limit $\underline{B}$ is a lower bound on the total radiosity function $B(x)$. □

3.2 Computing Upper Bounds

Unfortunately, we cannot use the same straightforward approach for computing the upper bounds $\overline{B} = (\overline{B}_1, \ldots, \overline{B}_n)^{\mathrm{T}}$. Let $\overline{K}$ be defined in the same way as $\underline{K}$, except that suprema are taken instead of infima. Then the matrix $I - \overline{K}$ is not necessarily diagonally dominant: the entries of $\overline{K}$ are upper bounds on the form-factors and thus the sum of the absolute values of the entries in a row of $I - \overline{K}$ can exceed 1. In such a case, iterative solution methods such as Jacobi or Gauss-Seidel could diverge. Furthermore, it is possible for $I - \overline{K}$ to be singular.

We now describe an iterative algorithm that transforms $\overline{K}$ into a strictly diagonally dominant matrix $\overline{K}^*$, such that the solution to the linear system of equations

$$\left(I - \overline{K}^*\right)\overline{B} = \overline{E}$$

is a vector of upper bounds. The solution $\overline{B}$ is computed simultaneously with the transformation of $\overline{K}$ into $\overline{K}^*$.

The algorithm starts by forming the matrix $\overline{K}$. We then perform Jacobi iterations using a modified matrix: because the sum of the form-factors from a given element cannot exceed 1, we can zero out elements in each row of $\overline{K}$ until the form-factors corresponding to the remaining non-zero entries sum to at most 1. To ensure that the modified matrix still yields an upper bound on the current light bounce, we zero out those entries corresponding to the dimmest elements. Since the brightness ordering of the elements may change from one iteration to another, different entries on the matrix may be zeroed out in each iteration. However, as the iterates converge to the solution $\overline{B}$, the order of the elements becomes fixed, at which point we obtain $\overline{K}^*$.

More precisely, at each iteration we sort the current solution vector $\overline{B}^{(k)}$ in order of decreasing brightness, and permute the columns of the matrix accordingly. Each entry in the solution vector is updated by:

$$\overline{B}_i^{(k+1)} = \overline{E}_i + \overline{\rho}_i \left[\sum_{j=1}^{l} \overline{F}_{ij}\overline{B}_j^{(k)} + \left(1 - \sum_{j=1}^{l} \overline{F}_{ij}\right) \overline{B}_{l+1}^{(k)} \right], \qquad (3)$$

where l is the largest index in row i such that

$$\sum_{j=1}^{l} \overline{F}_{ij} \leq 1.$$

It can be shown by induction that, if the iterates are sorted as described above, $\overline{B}^{(k)}$ is an upper bound on the radiosity after k interreflections. Thus, if the iterates converge, they converge to a total upper bound on the radiosity function $B(x)$.

Proof. To prove convergence, we must examine the behavior of each entry in the solution vector as the algorithm progresses. Because all the entries in the matrix $\overline{K}$ are non-negative, if $\overline{B}^{(0)}$ is set to $\overline{E}$, each entry in the solution vector increases at each iteration, i.e., for all i, $\overline{B}_i^{(k)} \leq \overline{B}_i^{(k+1)}$. Because we do not allow the form-factors in a row to exceed 1, each entry of $\overline{B}^{(k)}$ is bounded from above:

$$\overline{B}_i^{(k)} \leq E_{\max}\left(1 + \rho_{\max} + \cdots + \rho_{\max}^k\right) < \frac{E_{\max}}{1 - \rho_{\max}},$$

where $E_{\max}$ is the maximum emission and $\rho_{\max}$ the maximum reflectivity in the environment. From elementary analysis, we know that a monotonically increasing series that is bounded from above must converge. Since the vectors $\overline{B}^{(k)}$ are of a finite dimension, it follows that the vector sequence converges. □

```
GatherLowerBounds(node, lower)
foreach link ∈ node.links do
    lower += node.rho * link.Fmin * link.source.lower
end for
if IsLeaf(node) then
    node.lower = lower + node.emission
else
    node.lower = ∞
    foreach child ∈ node.children do
        GatherLowerBounds(child, lower)
        node.lower = min (node.lower, child.lower)
    end for
end if

GatherUpperBounds(node, contribList)
foreach link ∈ node.links do
    add the pair (link.Fmax, link.source.upper) to contribList
end for
if IsLeaf(node) then
    ffSum = 0
    node.upper = node.emission
    Sort(contribList)
    foreach pair (ff, upper) ∈ contribList do
        if ffSum + ff < 1 then
            ffSum += ff
            node.upper += node.rho * ff * upper
        else
            node.upper += node.rho * (1 − ffSum) * upper
            break
        end if
    end for
else
    node.upper = 0
    foreach child ∈ node.children do
        GatherUpperBounds(child, contribList)
        node.upper = max (node.upper, child.upper)
    end for
end if
```

Figure 1: *Pseudocode for gathering lower and upper bounds*

3.3 A Radiosity-Bounding Algorithm for HR

The algorithms in Sections 3.1 and 3.2 are presented in the context of the full matrix radiosity algorithm. They involve $O(n^2)$ work to compute the bounds on the form-factors, and therefore are not practical. However, they can be modified to work more efficiently within the HR framework. This requires a few changes to the standard HR algorithm. We need to compute and store upper and lower bounds *Fmin* and *Fmax* on the form-factor associated with each link. We can then gather, push, and pull radiosity bounds, in a way similar to that of the patch radiosities themselves. See Figure 1.

Note that gathering, pushing, and pulling are all performed in a single sweep of the hierarchy. Thus, the bounds are updated in place, which makes this a Gauss-Seidel iteration. This iteration converges faster than the Jacobi iteration used in previous HR algorithms. The speedup applies not only to the computation of bounds, but to the computation of the radiosities as well.

Gathering upper bounds is a bit more involved than gathering lower bounds, as can be seen from the pseudocode in Figure 1. *GatherUpperBounds* essentially pushes links down to the leaves, where they are sorted according to the brightnesses of the corresponding source nodes. The upper bound of the leaf node is then updated using equation (3).

One undesirable feature of this algorithm is the sorting of the links, which implies that $O(k \log k)$ work must be done to update the upper bound at each leaf node, where k is the number of the contributing links. However, complete sorting can be avoided in most cases. Sorting is only necessary when the sum of the contributing form-factors exceeds 1. Even then, only a few of the dimmest sources must be removed from the list. This can be done efficiently using a heap data structure with a *DeleteMin* operation. Tarjan [20] provides a detailed description of such data structures.

3.4 Bounding Form-Factors

Thus far we have assumed that the bounds $\underline{F}_{ij}$ and $\overline{F}_{ij}$ are available to us. How do we obtain them? We know of no analytical way of computing such bounds. Thus, we must resort to numerical optimization to find the minimum and the maximum values of the point-to-polygon form-factor from points on S_i to S_j. Such numerical methods are discussed in detail by Campbell [7]. These methods generally require the computation of form-factor gradients in the presence of occluders, which can be done either using finite differences or analytically, as described by Arvo [2].

In our current implementation we estimate form-factor bounds by evaluating the analytical point-to-polygon form-factor [18] at the center and at the vertices of each receiving patch. In the partially occluded case, the analytical formula is applied only to the visible parts of the source [16]. The upper and lower bounds on the form-factor are then set to the maximum and the minimum of these values. While these bounds are not conservative, as elements decrease in size the point-to-area form-factor function becomes monotonic over most elements, and the heuristic yields accurate bounds. This is particularly so when a discontinuity-driven subdivision strategy [15] is employed.

3.5 Results

The plots in Figures 2, 3, and 4 show the piecewise-constant bounding functions computed by the hierarchical bounding algorithm for three simple environments shown in Figure 5. The exact radiosity function is also plotted for each of the three cases. The functions are plotted along the dotted lines across the floor of each environment.

The bounds in Figure 2 were found in a single iteration, since there are no interreflections in this environment, and errors do not propagate. Because of this, the bounds are very tight: in fact, tighter piecewise constant bounds are only possible if more elements are used.

Figures 3 and 4 show two cases with interreflections, which results in looser bounds due to error propagation. As illustrated in the figures, the bounds computed by the algorithm become tighter as the accuracy of the solution increases. Note that the radiosity function and the bounds are not symmetric, because in the corresponding environments the reflectivity of the left wall is higher than that of the right wall.

4 COMPUTING REALISTIC ERROR ESTIMATES

Having obtained bounds on the radiosity of each element, we can obtain an upper bound on the local error there. Assuming that the

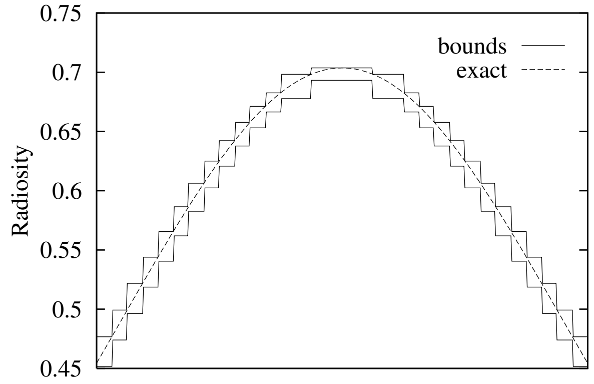

Figure 2: *Upper and lower bounds for the configuration shown in Figure 5a. The smooth curve in the middle is the exact radiosity function.*

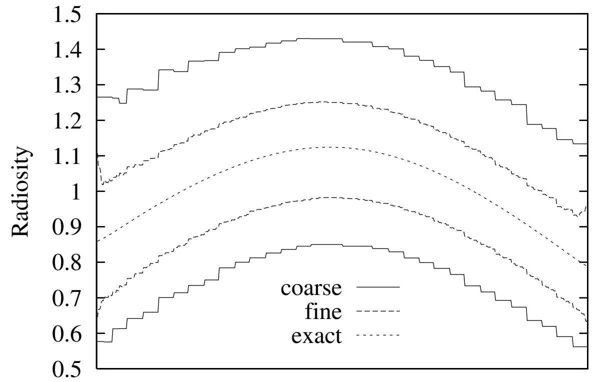

Figure 3: *Nested lower and upper bounds corresponding to two HR solutions for the configuration in Figure 5b. One solution is coarse and the other is fine. The smooth curve in the middle is the exact radiosity function.*

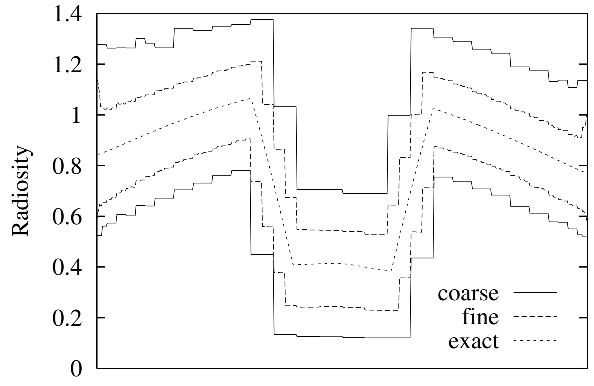

Figure 4: *Nested lower and upper bounds corresponding to two HR solutions for the configuration in Figure 5c. One solution is coarse and the other is fine. The smooth curve in the middle is the exact radiosity function.*

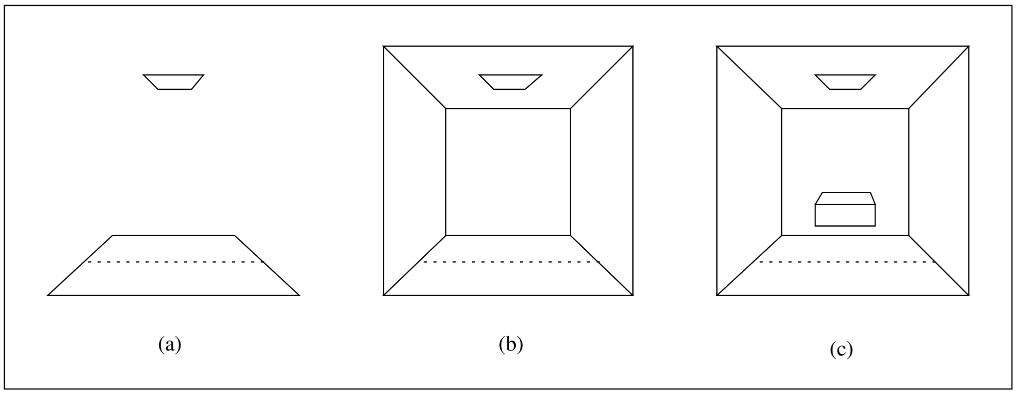

Figure 5: *Three simple test cases: (a) direct illumination only; (b) inter-reflections without occlusion; (c) inter-reflections with occlusion. All the surfaces are grey with reflectivities between 0.3 and 0.7.*

approximate radiosity B_i lies halfway between $\underline{B}_i$ and $\overline{B}_i$, the errors over S_i in the L_∞, L_1, and L_2 norms are:

$$\|B(x) - B_i\|_\infty \equiv \max_{S_i} |B(x) - B_i| \leq (\overline{B}_i - \underline{B}_i)/2, \quad (4)$$

$$\|B(x) - B_i\|_1 \equiv \int_{S_i} |B(x) - B_i|\, dx \leq A_i(\overline{B}_i - \underline{B}_i)/2, \quad (5)$$

$$\|B(x) - B_i\|_2^2 \equiv \int_{S_i} |B(x) - B_i|^2 dx \leq A_i(\overline{B}_i - \underline{B}_i)^2/4. \quad (6)$$

However, these error bounds represent a worst case scenario, and do not give an accurate error estimate even when the bounds on radiosity are as tight as in Figure 2. More realistic estimates are obtained if we assume that the values of $B(x)$ over S_i are uniformly distributed between $\underline{B}_i$ and $\overline{B}_i$. Under this assumption the L_1 and L_2 norm errors are estimated as:

$$\|B(x) - B_i\|_1 \approx A_i(\overline{B}_i - \underline{B}_i)/4, \quad (7)$$

$$\|B(x) - B_i\|_2^2 \approx A_i(\overline{B}_i - \underline{B}_i)^2/12. \quad (8)$$

We compute local error estimates at all the leaves of the hierarchy. These estimates are then pulled up to the roots of the hierarchies, and combined to obtain the estimate for the total error. The error at a parent node is the sum of the children's errors in the L_1 norm, and the maximum over the children's errors in L_∞ norm. For the L_2 norm, the squares of the children's errors are summed and the square root of the sum is taken.

4.1 Obtaining Tighter Non-Conservative Bounds

The conservative algorithm produces bounds which are quite loose even for simple environments, resulting in pessimistic error estimates. For most practical purposes good non-conservative error estimates could prove more useful than conservative error bounds. To quote Delves and Mohamed [10]:

> Given two error estimates of equal realism and equal cost and given that one is also a bound, we would clearly prefer the bound. In practice bounds usually prove expensive and pessimistic. Most users will then usually prefer a cheaper and more realistic error estimate, recognizing and accepting that estimates can sometimes err on the side of optimism.

We now present a way for computing tighter, although not conservative, bounds on radiosity. Error estimates computed from these radiosity bounds are not guaranteed error bounds, but they are much closer to the true errors. To compute these tighter bounds we approximate $\underline{B}$ and $\overline{B}$ as follows:

$$\underline{B} \approx \underline{E} + \underline{K}\widehat{B}$$

$$\overline{B} \approx \overline{E} + \overline{K}\widehat{B}.$$

This is equivalent to first setting both $\underline{B}$ and $\overline{B}$ to $\widehat{B}$ and then performing a single gathering iteration over the environment using the same *GatherLowerBounds* and *GatherUpperBounds* routines as before. Because only one iteration is performed, error propagation is no longer accounted for, but the computation is significantly simplified. More conservative estimates can be obtained by using two or more iterations, yielding a trade-off between speed and conservativity.

Two additional changes to the algorithm have been used to produce even tighter bounds. First, instead of bounding each form-factor separately and summing the bounds, we directly bound the radiosity contributed to each node through all its links: the contribution is evaluated at the center and at the vertices of the patch, and the extrema values are used as bounds. Second, we set *node.lower* and *node.upper* to the area-weighted averages of the children bounds, rather than to the minimum and the maximum, respectively.

4.2 Results

To test the quality of our error estimates we computed an extremely fine radiosity solution for each of the three environments in Figure 5. The floor polygon in each environment was sampled on a 400 by 400 grid to serve as a reference. We then computed a series of radiosity solutions of various degrees of accuracy for each environment. The floor polygon was sampled as before, and the difference between each solution and the reference solution was computed. This difference is referred to as the *measured* error. During each solution, two error estimates were computed: one from the conservative bounds, and the other from the tighter non-conservative bounds. In both cases equations (4), (7), and (8) were used to estimate the errors.

Figure 6 shows plots of the error estimates together with the measured error for each of the three test cases. The left column corresponds to L_∞ norm and the right column to L_1 norm. Error plots for L_2 norm are not given, since they look almost exactly the same as the L_1 plots for these test cases.

In the top pair of plots (corresponding to Figure 5a) the conservative and the non-conservative bounding algorithms yield the same bounds, and therefore the corresponding error estimates are always the same. For the L_∞ norm these estimates follow tightly the measured error, and for the L_1 norm they give realistic upper bounds.

In the remaining two cases there is a significant difference between the two types of error estimates. In the L_∞ norm, the non-conservative estimate does not always yield a bound on the error (see lower left plot). In the L_1 norm, both estimates bound the measured error from above, but the conservative error estimates are much more pessimistic than their non-conservative counterparts.

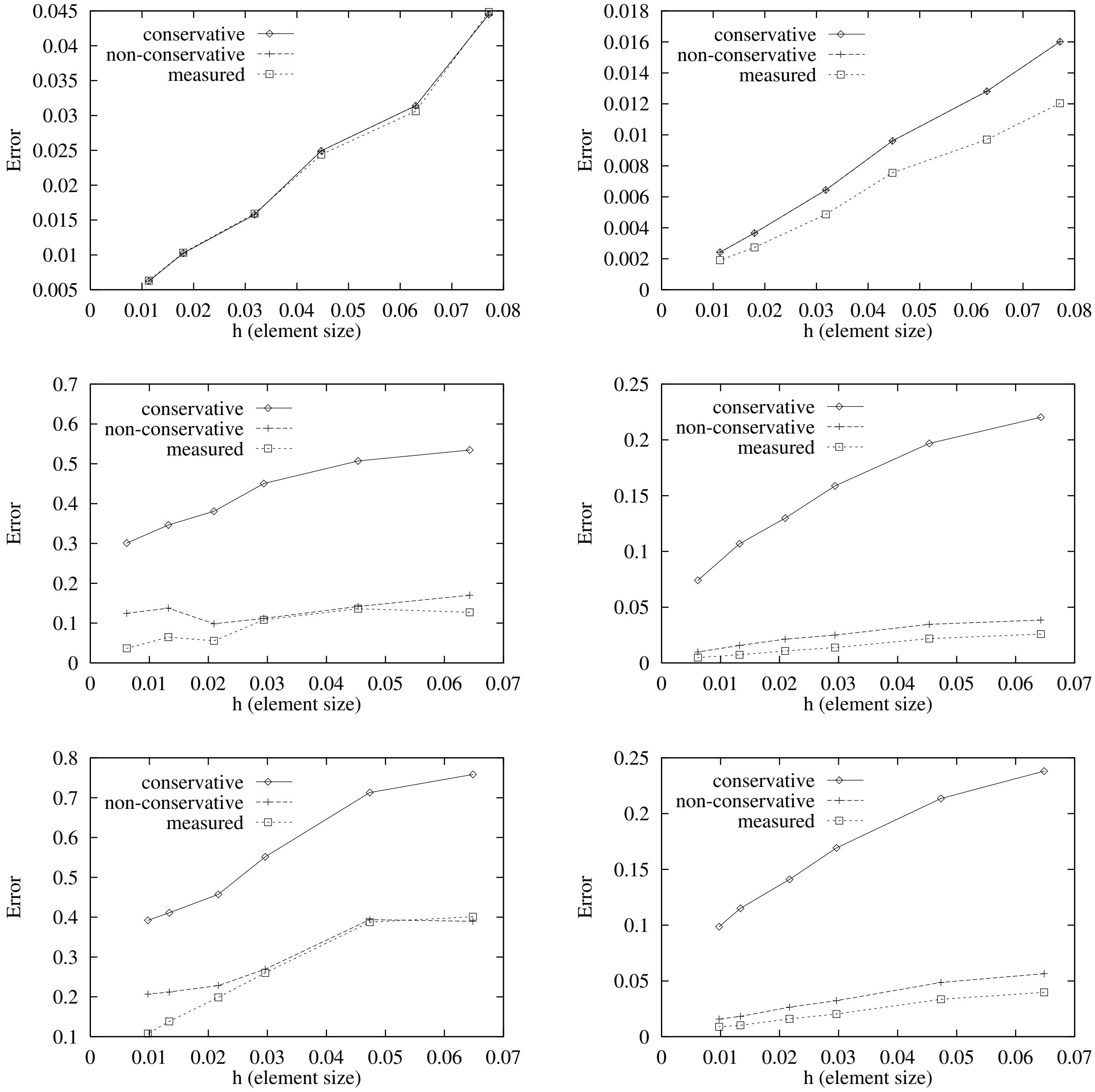

Figure 6: *Estimated and measured errors as a function of the element size for the three environments shown in Figure 5. The top row corresponds to Figure 5a, the middle row to 5b, and the bottom row to 5c. L_∞ errors are shown on the left and L_1 errors on the right.*

Both types of estimates become closer to the measured error as the accuracy of the solutions increases. Both the estimated and the measured errors vary linearly in the element size *h*, as expected for piecewise-constant approximations.

5 ERROR-DRIVEN REFINEMENT

The original hierarchical radiosity algorithm [13] adaptively improves the accuracy of a solution by estimating the error associated with each link. All links for which this error estimate exceeds a given threshold ϵ are refined and the solution is recomputed. More specifically, Hanrahan et al. suggest to refine all those links for which the estimated upper bound on the transferred energy exceeds ϵ. A link is refined by subdividing the node which has the greater form-factor, as seen from the center of the other node. This strategy is referred to as *brightness-weighted* refinement, as it gives priority to links transferring energy from bright sources.

We make two observations with respect to brightness-weighted refinement: (i) The links carrying the largest amount of energy are not necessarily the ones with the greatest errors in the transferred energy. (ii) Due to propagation of errors, refining the link between nodes *i* and *j* reduces not only the error over patch *i*, but may also reduce errors on other patches receiving light from it. Thus, to make the most benefit of each link refinement, priority should be given to links with the greatest effect on the total solution error.

The first observation suggests using $\Delta K_{ij} = \overline{K}_{ij} - \underline{K}_{ij}$ as the estimate of the link form-factor error, rather than using an upper bound on the form-factor. The error in the transferred energy is then given by $\Delta K_{ij}\overline{B}_j$.

To address the second observation, we derive an expression that relates ΔK_{ij} to the overall value of $\|\overline{B} - \underline{B}\|_1$, which is an upper bound on the overall error in L_1 norm. This is done by using the concept of importance.

5.1 Importance and Linear Functionals

Recall that *importance*, which we shall denote by Z, is the solution to the adjoint of the transport equation [14, 17]:

$$M^{\mathrm{T}} Z = R$$

where M^{T} is the adjoint of the real matrix M. In the context of radiosity, $M = I - K$. R is the *receiver* vector, which is dual to the emission vector E.

Any linear functional $f(B)$ of the radiosities can be expressed as an inner product $R^{\mathrm{T}}B$, where R_i gives the contribution of B_i to the value of $f(B)$. It can be shown [17] that

$$R^{\mathrm{T}} B = Z^{\mathrm{T}} E.$$

Thus, the i-th element of the importance vector gives the contribution made by a unit of emission at patch i to the value of $R^{\mathrm{T}}B$.

Note that because all the entries in the vector $\overline{B} - \underline{B}$ are non-negative, the value of its L_1 norm is given by a linear functional:

$$\left\|\overline{B} - \underline{B}\right\|_1 = R^{\mathrm{T}}\left(\overline{B} - \underline{B}\right),$$

where $R = [A_1, \ldots, A_n]^{\mathrm{T}}$ is the vector of patch areas. We can now derive the main result of this section:

$$\left\|\overline{B} - \underline{B}\right\|_1 = \underline{Z}^{\mathrm{T}}\Delta K\overline{B} + \underline{Z}^{\mathrm{T}}\left(\overline{E} - \underline{E}\right). \tag{9}$$

This expression relates $\|\overline{B} - \underline{B}\|_1$ to the link errors ΔK and to errors in the emission. The term $\underline{Z}^{\mathrm{T}}\Delta K\overline{B}$ is a sum over all links. The values summed are the link form-factor error ΔK_{ij} weighted by the lower bound on the importance of the receiver $\underline{Z}_i$ and the upper bound on the radiosity of the source $\overline{B}_j$. Intuitively, $\Delta K_{ij}\overline{B}_j$ is an upper bound on the incorrect radiosity that patch i receives from patch j, while $\underline{Z}_i$ gives the contribution of this incorrect radiosity on the total error bound. The $\underline{Z}^{\mathrm{T}}(\overline{E} - \underline{E})$ term gives the effect non-constant emission has on the total error bound.

Proof. Let $\overline{M} = I - \overline{K}^*$ and $\underline{M} = I - \underline{K}$. The system for lower bounds on the radiosities and its adjoint equation can now be written as

$$\underline{MB} = \underline{E} \quad \text{and} \quad \underline{M}^{\mathrm{T}}\underline{Z} = R,$$

and the upper bounds system can be written as $\overline{MB} = \overline{E}$. Noting that $\Delta K = \overline{K} - \underline{K} = \underline{M} - \overline{M}$, we obtain

$$\overline{MB} = (\underline{M} - \Delta K)\overline{B} = \overline{E},$$

from which it follows that

$$\underline{M}\overline{B} = \overline{E} + \Delta K\overline{B}. \tag{10}$$

Since $R^{\mathrm{T}}\underline{B} = \underline{Z}^{\mathrm{T}}\underline{E}$, we can express $R^{\mathrm{T}}(\overline{B} - \underline{B})$ in terms of ΔK as follows:

$$\begin{aligned} R^{\mathrm{T}}(\overline{B} - \underline{B}) &= R^{\mathrm{T}}\overline{B} - R^{\mathrm{T}}\underline{B} \\ &= R^{\mathrm{T}}\overline{B} - \underline{Z}^{\mathrm{T}}\underline{E} \\ &= (\underline{M}^{\mathrm{T}}\underline{Z})^{\mathrm{T}}\overline{B} - \underline{Z}^{\mathrm{T}}\underline{E} \\ &= \underline{Z}^{\mathrm{T}}\underline{M}\overline{B} - \underline{Z}^{\mathrm{T}}\underline{E}. \end{aligned}$$

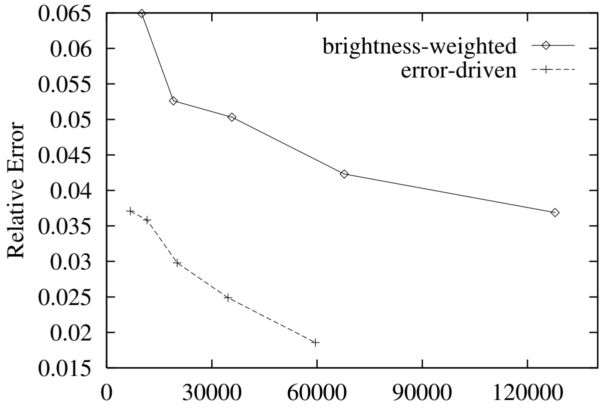

Figure 7: *Relative L_1 error as a function of the number of links: brightness-weighted vs. error-driven refinement.*

Substituting equation (10) into the above expression we have

$$\begin{aligned} R^{\mathrm{T}}(\overline{B} - \underline{B}) &= \underline{Z}^{\mathrm{T}}(\overline{E} + \Delta K\overline{B}) - \underline{Z}^{\mathrm{T}}\underline{E} \\ &= \underline{Z}^{\mathrm{T}}\Delta K\overline{B} + \underline{Z}^{\mathrm{T}}(\overline{E} - \underline{E}). \end{aligned}$$

□

Equation (9) is an exact expression relating errors in the transferred energy to the value of any linear functional $R^{\mathrm{T}}(\overline{B} - \underline{B})$. In particular this expression applies to the view-dependent functionals used by Smits et al. [17]. Thus, the expression that Smits et al. used to assess the view-dependent error can be viewed as an approximation to our more conservative bound.

5.2 A New Refinement Strategy

Equation (9) suggests a new error-driven refinement strategy for HR. In this strategy we refine all links for which $\underline{Z}_i\Delta K_{ij}\overline{B}_j > \epsilon$, since these links have the greatest effect on the total error bound. In order to evaluate this expression for each link, the routine *GatherLowerBounds* must be modified to solve for the lower bounds on importance in addition to lower bounds on radiosity. These changes are straightforward, and they do not significantly increase the amount of work done in the routine.

When we refine a link between a receiver node i and a source node j, we need to decide which node to subdivide to obtain the greatest reduction in the error. We use the following heuristic: node i is subdivided if

$$\left(\overline{K}_{ij} - \underline{K}_{ij}\right)\overline{B}_j \geq \overline{K}_{ij}\left(\overline{B}_j - \underline{B}_j\right),$$

and node j is subdivided otherwise. Intuitively, this means that we subdivide the receiver if the error due to the non-constant form-factor is greater than the error due to non-constant radiosity on the source.

5.3 Results

We implemented both brightness-weighted and error-driven refinement strategies and tested them on the simple environment shown in Figure 5c. The plot in Figure 7 shows how the measured relative L_1 error on the floor decreases as a function of the number of links. This plot demonstrates clearly that error-driven refinement is much more efficient: it is capable of achieving much smaller errors using only a fraction of the links required by brightness-weighted refinement. Similar results were observed on several other test cases.

6 CONCLUSIONS

We have described a hierarchical radiosity algorithm that computes conservative bounds on the exact radiosity function simultaneously with the computation of the approximate solution. In contrast to previous attempts to estimate the errors in radiosity, our algorithm properly accounts for the propagation of errors due to interreflections, and provides conservative upper bounds on the error. We have also presented a non-conservative version of this algorithm that provides tighter bounds and more realistic error estimates. Finally, we have derived a new error-driven adaptive refinement strategy for hierarchical radiosity that significantly outperforms brightness-weighted refinement.

We hope that the results described in this paper will prove useful not only for reliable image synthesis, but also for other fields in which similar integral equations arise. Examples include radiative heat transfer, illumination engineering, and remote sensing. In all these fields numerical accuracy of the results is even more crucial than in image synthesis.

Much further work is needed on *a posteriori* error analysis for radiosity. For instance, we need techniques for efficiently computing reliable tight bounds on form-factors, perhaps using tools such as irradiance Jacobians [2]. Also, we would like to extend the techniques described in this paper to higher order basis functions. While we are able to bound the exact radiosity function, our bounds will not necessarily apply to higher order approximations, as they may exhibit oscillations [22]. Even if the approximation is contained within our bounds, equations (7) and (8) will probably produce pessimistic error estimates, since they assume that the approximation is piecewise-constant.

Prior to rendering an image, the radiosity solution is usually projected onto a new set of basis functions, typically piecewise-linear or piecewise-quadratic [15]. We need to analyze the effect of such projections on the error in order to extend our error estimates from solutions to images.

Finally, we would like to emphasize the need for the development of perceptual error metrics for image synthesis. Fast convergence in a quantitative error metric, such as the L_1 norm, does not necessarily imply fast convergence of the resulting images, as perceived by a human observer. Thus, we expect that a perceptual error-driven refinement strategy would be more appropriate for image synthesis.

ACKNOWLEDGEMENTS

Special thanks go to Jim Arvo: without his encouragement and advice this paper would not have been written. We would also like to thank Richard Lobb and Charlie Van Loan for helpful discussions and for reviewing the manuscript. The comments provided by the anonymous reviewers have helped greatly to improve this paper. This work was supported by the NSF grant, "Interactive Computer Graphics Input and Display Techniques" (CCR-8617880), by the NSF/DARPA Science and Technology Center for Computer Graphics and Scientific Visualization (ASC-8920219), and by generous donations of equipment from Hewlett-Packard.

REFERENCES

[1] ANSELONE, P. M. Convergence and Error Bounds for Approximate Solutions of Integral and Operator Equations. In *Error in Digital Computation*, L. B. Rall, Ed., vol. 2, John Wiley & Sons, New York, 1965, pp. 231–252.

[2] ARVO, JAMES. The Irradiance Jacobian for Partially Occluded Polyhedral Sources. In *Computer Graphics* Proceedings, Annual Conference Series, 1994.

[3] ARVO, JAMES, KENNETH TORRANCE, AND BRIAN SMITS. A Framework for the Analysis of Error in Global Illumination Algorithms. In *Computer Graphics* Proceedings, Annual Conference Series, 1994.

[4] BABUŠKA, I., O. C. ZIENKIEWICZ, J. GAGO, AND E. R. DE A. OLIVEIRA, EDS. *Accuracy Estimates and Adaptive Refinements in Finite Element Computations*, John Wiley & Sons, Chichester, 1986.

[5] BREBBIA, C. A. AND M. H. ALIABADI, EDS. *Adaptive Finite and Boundary Element Methods*, Computational Mechanics Publications, Southampton, and Elsevier Applied Science, London, 1993.

[6] BROWN, R. W. Upper and Lower Bounds for Solutions of Integral Equations. In *Error in Digital Computation*, L. B. Rall, Ed., vol. 2, John Wiley & Sons, New York, 1965, pp. 219–230.

[7] CAMPBELL, III, A. T. *Modeling Global Diffuse Illumination for Image Synthesis*, PhD dissertation, University of Texas at Austin, Austin, Texas, December 1991.

[8] COHEN, MICHAEL F., DONALD P. GREENBERG, DAVID S. IMMEL, AND PHILIP J. BROCK. An Efficient Radiosity Approach for Realistic Image Synthesis. *IEEE Computer Graphics and Applications*, 6(2), March 1986, pp. 26–35.

[9] COHEN, MICHAEL F. AND JOHN R. WALLACE. *Radiosity and Realistic Image Synthesis*, Academic Press Professional, Cambridge, Massachusets, 1993.

[10] DELVES, L. M. AND J. L. MOHAMED. *Computational Methods for Integral Equations*, Cambridge University Press, Cambridge, Great Britain, 1985.

[11] GOLUB, GENE H. AND CHARLES F. VAN LOAN. *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 2nd edition, 1989.

[12] GORTLER, STEVEN J., PETER SCHRÖDER, MICHAEL F. COHEN, AND PAT HANRAHAN. Wavelet Radiosity. In *Computer Graphics* Proceedings, Annual Conference Series, 1993, pp. 221–230.

[13] HANRAHAN, PAT, DAVID SALZMAN, AND LARRY AUPPERLE. A Rapid Hierarchical Radiosity Algorithm. *Computer Graphics*, 25(4), July 1991, pp. 197–206.

[14] LEWINS, JEFFERY. *Importance, The Adjoint Function: The Physical Basis of Variational and Perturbation Theory in Transport and Diffusion Problems*, Pergamon Press, New York, 1965.

[15] LISCHINSKI, DANI, FILIPPO TAMPIERI, AND DONALD P. GREENBERG. Combining Hierarchical Radiosity and Discontinuity Meshing. In *Computer Graphics* Proceedings, Annual Conference Series, 1993, pp. 199–208.

[16] NISHITA, TOMOYUKI AND EIHACHIRO NAKAMAE. Half-Tone Representation of 3-D Objects Illuminated by Area Sources or Polyhedron Sources. Proceedings of COMPSAC '83 (Chicago, Illinois, November 1983), pp. 237–241.

[17] SMITS, BRIAN E., JAMES R. ARVO, AND DAVID H. SALESIN. An Importance-Driven Radiosity Algorithm. *Computer Graphics*, 26(4), July 1992, pp. 273–282.

[18] SPARROW, E. M. A New and Simpler Formulation for Radiative Angle Factors. *ASME Journal of Heat Transfer*, 85(2), May 1963, pp. 81–88.

[19] SZABÓ, BARNA AND IVO BABUŠKA. *Finite Element Analysis*, John Wiley & Sons, New York, 1991.

[20] TARJAN, ROBERT ENDRE. *Data Structures and Network Algorithms*, SIAM, Philadelphia, 1983.

[21] TROUTMAN, ROY AND NELSON L. MAX. Radiosity Algorithms Using Higher Order Finite Elements. In *Computer Graphics* Proceedings, Annual Conference Series, 1993, pp. 209–212.

[22] ZATZ, HAROLD R. Galerkin Radiosity: A Higher Order Solution Method for Global Illumination. In *Computer Graphics* Proceedings, Annual Conference Series, 1993, pp. 213–220.