Hierarchical Solution Techniques for Realistic Rendering

François X. Sillion

HAL Id: inria-00510129

https://hal.inria.fr/inria-00510129

Submitted on 17 Aug 2010

Hierarchical Solution Techniques for Realistic Rendering

François Sillion *

CNRS
*i*MAGIS
B.P. 53, 38041 Grenoble Cedex 9, France.

Abstract

The computation of realistic images requires the accurate simulation of radiant energy exchanges in a scene. The algorithms used to conduct such simulations are usually not restricted to visible light, thus the development of efficient computational techniques is of interest for a wide variety of applications (infrared and radar imaging, radio propagation etc.). Most current realistic rendering algorithms are based either on a deterministic, finite-element type approach (for instance, the radiosity method) or on a stochastic process (Monte Carlo methods). Recent advances in this field have promoted the use of hierarchical data structures and algorithms to represent radiant exchanges at different resolutions. The ability to always select the level of representation most appropriate for each phase of the simulation yields an impressive reduction of the total computational expense for a given accuracy level. Significantly, it also creates truly progressive algorithms that can produce an approximate solution in a very short time, and continuously improve it over time.

This documents summarizes a STAR lecture, and presents several hierarchical algorithms in the context of the radiosity method. The most recent clustering techniques, allowing the simulation of very complex scenes, are reviewed. Pending research issues are identified and discussed, in particular the difficult question of error control for hierarchical methods.

1 Rendering and lighting simulation

In the history of computer graphics, the notion of *rendering* has evolved considerably. In the early days, rendering was about creating a picture of a three-dimensional scene, and dealt primarily with the appearance of each surface, in terms of shading, highlights, transparency and texture. More recently, the quest for realism prompted a new interest in *lighting simulation*, whereby the global aspects of illumination are investigated: interreflection effects, in particular, play a key role in the general atmosphere of a scene, providing subtle but decisive cues about the threedimensional arrangement of objects.

Interestingly, lighting simulation techniques are now spreading outside research laboratories to become real tools, with practical and industrial applications. As we shall see, two conditions for the effective application of these techniques on "real-world" problems are (a) controlling the quality of the simulation, and (b) harnessing the algorithmic complexity of the calculation.

In this presentation, we are concerned with the simplified problem of lighting simulation for a set of diffuse surfaces. The application to image synthesis is obvious, but applications for other spectral regions are possible, such as thermal infrared, radar, or centimeter waves.

In the case of ideal diffuse surfaces, the balance of radiant energy

exchanges in a scene is represented by an integral equation, expressing the *radiosity*[1] at each point:

$$B(x) = E(x) + \rho_d(x) \int_{y \in S} B(y) \frac{\cos\theta \cos\theta'}{\pi r^2} V(x,y) dy. \quad (1)$$

Notations are illustrated in Figure 1 and will not be further discussed. We shall now focus on particular algorithms, and refer the reader to general books for background material [5, 20].
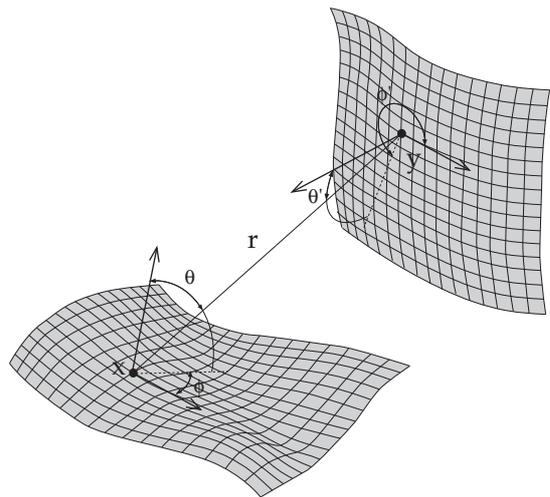

Figure 1: Geometry of radiant energy exchange.

1.1 The radiosity algorithm

The radiosity method consists of transforming Equation 1 into a linear system of equations of finite size. This can be done by projecting the equation onto a set of basis functions, and looking for the best possible approximation of the solution in the space spanned by these functions. The original radiosity method, for instance, uses a number of surface patches with the assumption of a constant radiosity function across each patch. Coupling coefficients between these patches, called *form factors*, are evaluated using a variety of numerical methods. The resolution of the linear system of equations then yields a *radiosity solution*, that is, a set of patch radiosity values. This solution is independent of the viewing conditions, and therefore particularly suited for lighting applications: it is possible to move into the virtual scene at a marginal cost, since the lighting calculation is only performed once.

---

[1]Total (emitted and reflected) flux density radiated at the point.

1.2 Limitations of the plain radiosity method

The main drawback of the radiosity approach is the number of form factors, quadratic in the number of surface patches used. The accuracy of the calculation is directly linked to the minimal size of the surface elements, but dividing the side length of all elements by a factor of 2 results in a 16-fold increase in the number of form factors! Therefore it is impossible in practice to simulate scenes with more than a few thousand surfaces to a reasonable accuracy, by storing a complete matrix of form factors.

A progressive formulation of the radiosity algorithm has been proposed, allowing the visualization of partial results, and avoiding the storage of a full form factor matrix [3, 6]. However the asymptotic complexity for a converged solution is not improved.

Most practical implementations of radiosity for image synthesis applications make use of adaptive meshing techniques, in an attempt to control the growth of the number of surface patches for a given accuracy [4]. Figure 2 shows an example where smaller elements have been generated only in the vicinity of a shadow boundary.
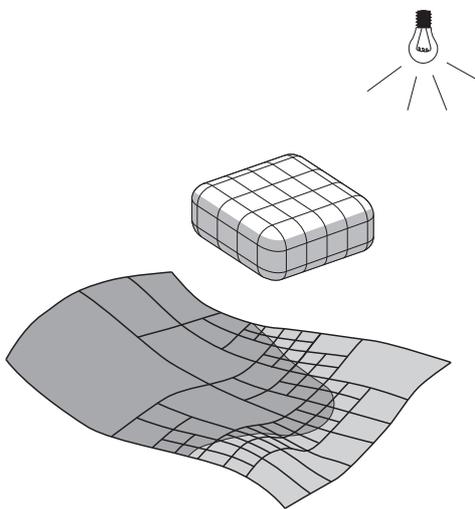

Figure 2: Adaptive subdivision generated along a shadow boundary.

Note also that the computation of form factors entails the evaluation of visibility between pairs of surfaces, a potentially very costly operation. This cost can either be incurred in computation time, or in required memory space, when specific visibility data structures are used [23].

2 Hierarchical formulation of radiosity

The recent introduction of a hierarchical reformulation of the radiosity algorithm has partially solved the dilemma of choosing between accuracy and speed. Its principle is quite simple, and extends the notion of adaptive subdivision to all pairwise interactions.

The modified algorithm uses a hierarchy of representations for a given surface, corresponding to different levels of detail. Let us first discuss the case of constant radiosity patches. In that case, a possible data structure to hold the multi-resolution radiosity information is the *quadtree*, for quadrilateral elements.

When considering the exchange of energy between two surfaces, the appropriate level of detail can be selected on each surface, allowing to meet the accuracy requirements without wasting resources. In other words, a given surface may be finely subdivided for the accurate simulation of its interaction with nearby objects; however, each of the resulting elements is not necessarily interacting with the entire scene. Instead, exchanges with more distant surfaces are treated at a higher, simplified level [9].

2.1 Form factors and links

The hierarchical radiosity algorithm represents energy exchanges using a network of *links* between surface patches. Each link represents a radiant interaction and conceptually carries some energy. It can be associated with a block in the form factor matrix between the elements of minimal size, as shown in Figure 3.
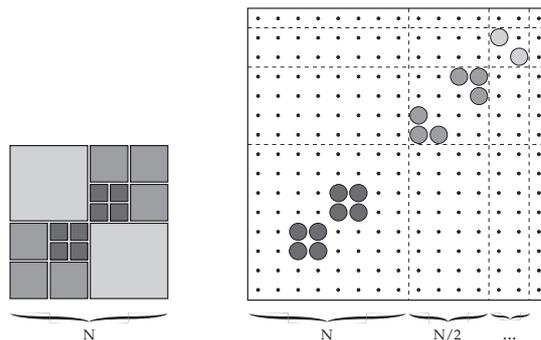

Figure 3: Form factor matrix for a one-dimensional example, where N is the number of elements at the finest level of subdivision. (left) each block corresponds to a link, thus large block represent links between large surface patches, composed of several terminal elements. (right) a view of the form factor matrix for all patches in the hierarchy, including non-terminal patches.

Links are created in a top-down recursive refinement procedure. For each pair of input surfaces, a link is first tentatively created, and passed on to the refinement procedure. The goal of this procedure is to decide whether the interaction can be correctly represented at the current level. A variety of criteria can be used to make this decision, some of which are reviewed in the next section. In cases where the current level is deemed unsatisfactory, one of the patches is subdivided, and four new links (in the case of a quadtree structure) are created with its children, and passed to the refinement procedure. The refinement procedure terminates when all links are "accurate enough" or connect surfaces smaller than a threshold.

An example showing the links generated between a pair of orthogonal polygons is presented in Figure 4. In this particular case the situation of adjacent polygons is independent of scale, therefore the subdivision was halted by the minimum area threshold. Note that each patch in the hierarchy ends up being linked to the same number of other patches. This property is quite general and results in a total number of links linear in the total number of surface elements created. In other words, going back to Figure 3, the number of entries computed in the right-hand matrix (or, equivalently, the number of blocks in the left-hand matrix) is $\mathcal{O}(N)$.

The hierarchical radiosity algorithm proceeds by first refining the set of links, then propagating energy across those links, and finally maintaining the consistency of the representation throughout the hierarchy. The propagation step simply requires traversing the structure and computing the exchange of radiosity across each link, using the associated form factor. The consistency step, also called the *Push/Pull* procedure, ensures that all elements of the hierarchy have a complete view of the energy physically received, including that received at higher levels of the hierarchy and that received at lower levels.
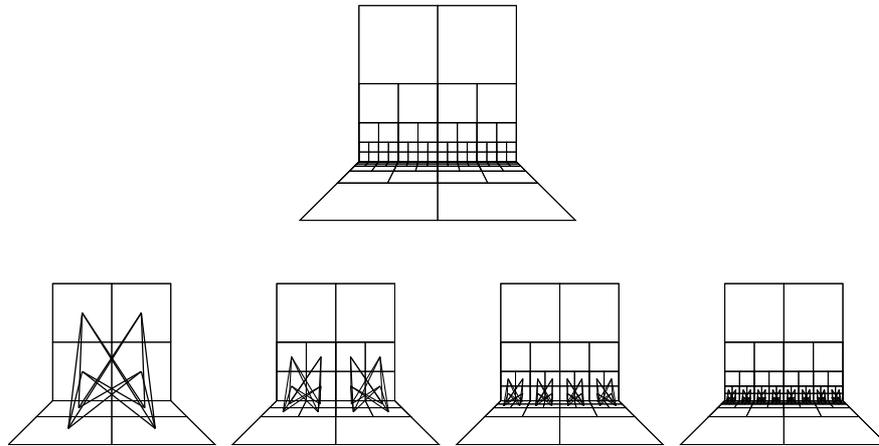
Figure 4: Hierarchical subdivision obtained for two perpendicular polygons *(Courtesy of P. Hanrahan, Princeton University)*

2.2 Refinement criteria

The core of the hierarchical radiosity algorithm is the link refinement procedure, which decides whether two patches can interact directly, or if it is better to push their interaction down the hierarchy. The subdivision criterion must determine whether the approximation incurred by using the current level is acceptable.

Criteria that have been proposed range from a simple bound on the magnitude of the form factor [8] to the estimation of the impact of the link on a global measure of the solution error [11]. Using the notion of *importance* [22], the subdivision control process can incorporate view-dependent information, encouraging subdivision in areas that are visible or otherwise contribute to the final result. In general, more elaborate criteria allow a more accurate control of the subdivision, thereby saving the expense of unnecessary refinement [10]. On the other hand, care must be taken that they do not become to costly to evaluate.

Note that the computation of form factors must be done with a flexible algorithm, in order to allow both a crude, but fast, estimation for use in the subdivision criteria, and a more precise calculation at the time of link creation. For this reason ray-tracing based techniques are generally preferred.

2.3 Wavelets

Recent work has shown that the hierarchical radiosity algorithm is equivalent to the use of a simple wavelet basis for the resolution of Equation 1 using finite elements. The value of this angle of view is that it provides a natural avenue for the development of more efficient algorithms using other wavelet bases.

Several researchers have applied the wavelet formalism to the radiosity equation, both for diffuse [7, 14] and non-diffuse [2, 15, 12] surfaces. The use of higher-order basis functions allows a reduction of the number of elements, at the cost of a more expensive calculation for each coupling coefficient. Continuity issues require particular attention when dealing with non-constant basis functions.

2.4 Application to volumes

For the case of scattering volumes, the zonal method, which derives from the radiosity method, can also be reformulated in a hierarchical manner. Notable differences include the possibility of *self-links* connecting a volume to itself (to represent internal energy exchanges), and the likely inhomogeneous character of high-level volumes. This second aspect requires a slight modification of the push/pull procedure [17]. Figure 5 shows an example scene where surfaces and volumes are integrated in a single hierarchy.

Figure 5: Result of a hierarchical simulation for a scene combining surfaces and scattering volumes.

3 Clustering techniques

Analyzing the complexity of the hierarchical radiosity algorithm shows that the total number of links, while linear in the total number of hierarchical elements created, remains quadratic in the number of input surfaces (since each pair requires an initial link). Therefore it is very efficient for scenes where a small number of large surfaces exchange energy. However it does not allow the simulation of scenes containing many small objects, with little need for hierarchical subdivision. Unfortunately, this type of scene is frequently found in architectural simulation, as soon as furniture and objects of daily life are incorporated.

3.1 Building and using clusters

The idea of using a simplified representation of the radiant properties of a collection of surfaces was first applied in the context of Monte Carlo simulation [13]. Manually constructed clusters were assigned equivalent reflectance properties, and were used to solve a simplified global illumination problem. This intermediate solution was then used as the basis of an image-based Monte Carlo calculation, providing the secondary illumination component.

In the context of radiosity, early attempts separated the scene into sub-components, and later combined the corresponding solutions [24, 1]. In order to efficiently reduce the complexity of the radiosity algorithm in cases with many small objects, a *complete* hierarchy can be created, grouping nearby objects in *clusters* that can directly exchange energy[16, 21]. In order for the method to be aplicable to general scenes, it is important that the hierarchy of clusters be constructed automatically. Current algorithms employ either contents-based strategies (such as hierarchies of bounding volumes [18]) or fixed subdivision methods (octrees or k-d trees). The latter have several algorithmic advantages but typically generate clusters that are less suited to accurate error estimation.

Figure 6 illustrates the notion of cluster, and shows links from the selected cluster to other surfaces (in this case, light sources). By establishing links at the cluster level, it was possible to avoid considering the detailed cluster contents.

Figure 6: Example of a cluster, with links to other radiative elements.

3.2 Algorithmic complexity

Subdivision criteria for links between clusters can be more complex than between surfaces: for instance one possibility is to compute a bound on the transfered energy by enumerating each cluster's contents. For two clusters containing p and q objects, respectively, such an enumeration requires $\mathcal{O}(p + q)$ work, as opposed to the $\mathcal{O}(pq)$ work required in a naive radiosity approach.

Clustering effectively reduces the algorithmic complexity (in terms of the number of links) to $\mathcal{O}(n \log n)$ for n input surfaces, thereby paving the way for efficient simulation in very complex scenes. In particular, a continuous range of variation is now possible for the accuracy of the results. Crude images can be produced very quickly (in a matter of seconds, for scenes containing over 10000

surfaces) and refined over time. This capability is especially useful for design applications.

3.3 Multiresolution visibility

While clustering successfully reduces the number of links, each of these links still requires the calculation of visibility information, a costly operation in general.

When a hierarchy of clusters is available, it can be used to perform visibility calculations in a multi-resolution manner: shadows are only computed down to a given level of detail, as specified by the user [18]. This is done by exploiting an analogy between a cluster of surfaces and a semi-transparent volume [17].

Figure 7 shows an example of the multi-resolution visibility paradigm. The algorithm was run on a scene with the same energy-based error threshold each time, but with different levels of detail requested in the shadows. Note that shadow detail is effectively traded for speed.
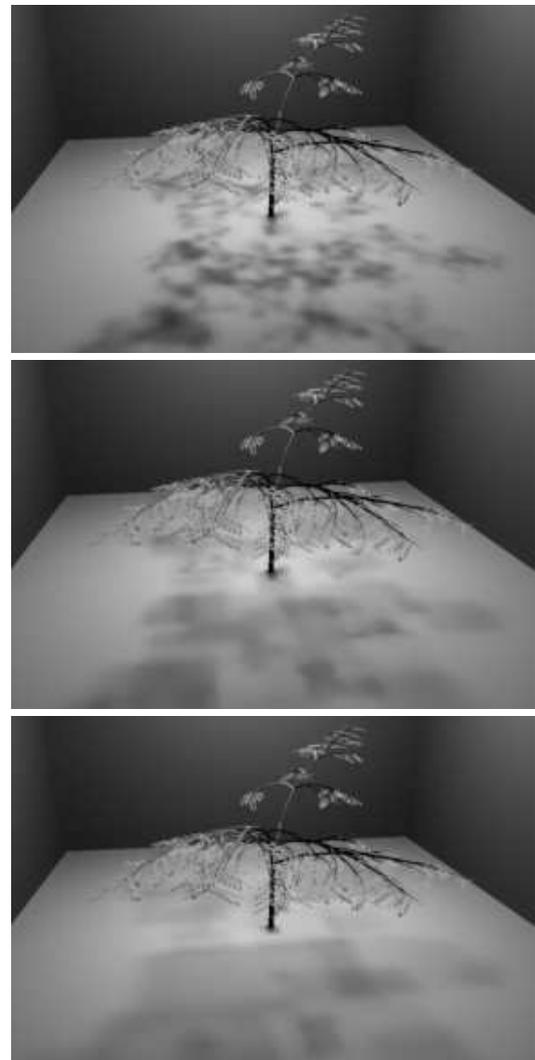

Figure 7: Increasing the desired feature size (Fsize) reduces both the amount of subdivision and the cost of visibility computations. (top) Fsize = 0, 621s. (middle) Fsize = 4, 245s. (bottom) Fsize = 8, 148s. Tree courtesy of CIRAD, modelled with 7967 polygons using AMAP.

3.4 Radiance calculation

Hierarchical radiosity can be expressed in a unified manner for diffuse surfaces and isotropic scattering volumes, as shown in [17]. However, the ideal case of isotropic scattering does not exist in practice. Similarly, when using object clusters, it cannot be safely assumed that they behave as isotropic scatterers. Even if all surfaces are diffuse, a cluster containing some of them will reflect light in a directional manner. This is the reason why all surfaces in a cluster should be considered individually when estimating the cluster's contribution to an energy transfer.

Clustering radiosity can also be extended to explicitly represent such directional dependencies. By storing a directional distribution of radiant intensity with each cluster, it is possible to estimate energy transfer simply by querying these distributions [19]. In addition, the same formalism allows the consistent treatment of non-diffuse surfaces. Figure 8 shows the directional distribution of radiant intensity for a cluster of surfaces.
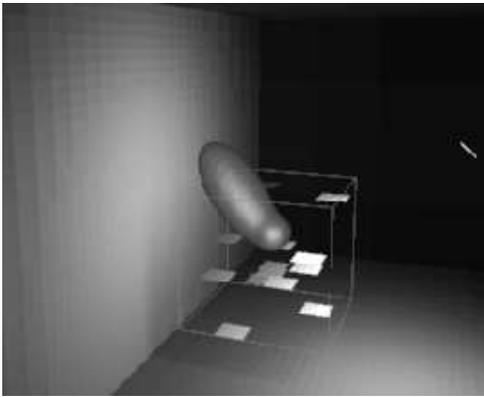

Figure 8: Plot of the directional distribution of radiant intensity for a cluster of glossy surfaces illuminated by a light source.

4 Conclusions and research directions

The introduction of hierarchical algorithms for quantitative lighting simulation has been a major step towards the actual diffusion and usage of simulation techniques. Applications for the research performed in this area go beyond the lighting problem since, for instance, clustering techniques can be very useful in modeling. Many algorithms need to be improved however, and research on this subject is likely to remain active in coming years. Examples of open problems are the analysis of the relative performance of the various subdivision criteria, the definition of better quantitative criteria based on error estimates, or the use of wavelets on triangular domains.

References

[1] Bruno Arnaldi, Xavier Pueyo, and Josep Vilaplana. On the division of environments by virtual walls for radiosity computation. In P. Brunet and F.W. Jansen, editors, *Photorealistic Rendering in Computer Graphics*, pages 198–205. Springer Verlag, 1993. Proceedings of the Second Eurographics Workshop on Rendering (Barcelona, Spain, May 1991).

[2] Per Christensen, Eric Stollnitz, David Salesin, and Tony DeRose. Wavelet radiance. In *Proceedings of Fifth Eurographics Workshop on Rendering*, Darmstadt, Germany, June 1994. Eurographics.

[3] Michael F. Cohen, Shenchang Eric Chen, John R. Wallace, and Donald P. Greenberg. A progressive refinement approach to fast radiosity image generation. *Computer Graphics*, 22(4):75–84, August 1988. Proceedings SIGGRAPH '88 in Atlanta, USA.

[4] Michael F. Cohen, Donald P. Greenberg, David S. Immel, and Philip J. Brock. An efficient radiosity approach for realistic image synthesis. *IEEE Computer Graphics and Applications*, 6(3):25–35, March 1986.

[5] Michael F. Cohen and John R. Wallace. *Radiosity and Realistic Image Synthesis*. Academic Press, Boston, 1993.

[6] Steven Gortler, Michael F. Cohen, and Philipp Slusallek. Radiosity and relaxation methods. *IEEE Computer Graphics and Applications*, 14(6):48–58, November 1994.

[7] Steven J. Gortler, Peter Schröder, Michael F. Cohen, and Pat Hanrahan. Wavelet radiosity. In *Computer Graphics Proceedings, Annual Conference Series:* SIGGRAPH '93 (Anaheim, CA, USA), pages 221–230. ACM SIGGRAPH, New York, August 1993.

[8] Pat Hanrahan and David Saltzman. A rapid hierarchical radiosity algorithm for unoccluded environments. In C. Bouville and K. Bouatouch, editors, *Photorealism in Computer Graphics*. Springer Verlag, EurographicSeminars series, 1992.

[9] Pat Hanrahan, David Saltzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. *Computer Graphics*, 25(4):197–206, August 1991. Proceedings SIGGRAPH '91 in Las Vegas (USA).

[10] Nicolas Holzschuch, François Sillion, and George Drettakis. An efficient progressive refinement strategy for hiearchical radiosity. In *Fifth Eurographics Workshop on Rendering*, Darmstadt, Germany, June 1994. In these proceedings.

[11] Dani Lischinski, Brian Smits, and Donald P. Greenberg. Bounds and error estimates for radiosity. In *Computer Graphics Proceedings, Annual Conference Series:* SIGGRAPH '94 (Orlando, FL). ACM SIGGRAPH, New York, July 1994.

[12] Sumant Pattanaik and Kadi Bouatouch. Haar wavelet: A solution to global illumination with general surface properties. In *Proceedings of Fifth Eurographics Workshop on Rendering*, Darmstadt, Germany, June 1994. Eurographics.

[13] Holly Rushmeier, Charles Patterson, and Aravindan Veerasamy. Geometric simplification for indirect illumination calculations. In *Proceedings Graphics Interface '93*. Morgan Kaufmann publishers, 1993.

[14] Peter Schröder, Steven J. Gortler, Michael F. Cohen, and Pat Hanrahan. Wavelet projections for radiosity. In *Proceedings of Fourth Eurographics Workshop on Rendering*, pages 105–114. Eurographics, June 1993. Technical Report EG 93 RW.

[15] Peter Schröder and Pat Hanrahan. Wavelet methods for radiance calculations. In *Proceedings of Fifth Eurographics Workshop on Rendering*, Darmstadt, Germany, June 1994. Eurographics.

[16] François Sillion. Clustering and volume scattering for hierarchical radiosity calculations. In *Proceedings of Fifth Eurographics Workshop on Rendering*, Darmstadt, Germany, June 1994. Eurographics.

[17] François Sillion. A unified hierarchical algorithm for global illumination with scattering volumes and object clusters. *to appear in IEEE Transactions on Visualization and Computer Graphics*, 1(3), September 1995. (a preliminary version appeared in the fifth Eurographics workshop on rendering, Darmstadt, Germany, June 1994).

[18] François Sillion and George Drettakis. Feature-based control of visibility error: A multiresolution clustering algorithm for global illumination. In *Computer Graphics Proceedings, Annual Conference Series:* SIGGRAPH '95 (Los Angeles, CA). ACM SIGGRAPH, New York, August 1995.

[19] François Sillion, George Drettakis, and Cyril Soler. A clustering algorithm for radiance calculation in general environments. In *Sixth Eurographics Workshop on Rendering*, Dublin, Ireland, June 1995.

[20] François Sillion and Claude Puech. *Radiosity and Global Illumination*. Morgan Kaufmann publishers, San Francisco, 1994.

[21] Brian Smits, James Arvo, and Donald P. Greenberg. A clustering algorithm for radiosity in complex environments. In *Computer Graphics Proceedings, Annual Conference Series:* SIGGRAPH '94 (Orlando, FL), pages 435–442. ACM SIGGRAPH, New York, July 1994.

[22] Brian E. Smits, James R. Arvo, and David H. Salesin. An importance-driven radiosity algorithm. *Computer Graphics*, 26(4):273–282, July 1992. Proceedings of SIGGRAPH '92 in Chicago (USA).

[23] Seth J. Teller and Patrick M. Hanrahan. Global visibility algorithms for illumination computations. In *Computer Graphics Proceedings, Annual Conference Series:* SIGGRAPH '93 (Anaheim, CA, USA), pages 239–246. ACM SIGGRAPH, New York, August 1993.

[24] H. Xu, Q. Peng, and Y Liang. Accelerated radiosity method for complex environments. *Eurographics '89*, pages 51–61, September 1989.