

PYRAMIDAL HEMISPHERE SUBDIVISION RADIOSITY. DEFINITION AND IMPROVEMENTS

Vincent Jolivet[†], Dimitri Plemenos[†] and Mateu Sbert[‡]

[†]Laboratoire MSI, Université de Limoges,
1, rue d'Isle, F-87000 Limoges, France.
e-mail: jolivet@unilim.fr, plemenos@unilim.fr
homepage URL: <http://msi.unilim.fr/~jolivet>, <http://msi.unilim.fr/~plemenos>

[‡]Institut d'Informàtica i Aplicacions, Universitat de Girona
Lluís Santaló s/n, E-17071 Girona, Spain.
e-mail: mateu@ima.ugd.es

ABSTRACT

In this paper a complete presentation of Pyramidal Hemisphere subdivision method is given. This method improves Monte Carlo radiosity by sending more rays towards selected directions. More precisely, we determine regions of the scene where the distribution of the energy must be done more accurately. The number of rays sent in a direction is function of a pyramidal region, defined by the center of the shooting patch and a spherical triangle on the surface of a hemisphere surrounding the patch, and the number of patches contained in this region. Thus, the rays shot from a patch have not all the same energy. This method allows us not only to obtain fine details much sooner and with lower cost, but also the overall efficiency is considerably increased. Next we present a new way to calculate the number of rays sent in a region by using a dynamic estimation of the visible patches in this region.

Keywords: Monte Carlo radiosity, hemisphere subdivision, heuristic search.

1. INTRODUCTION

Radiosity has been introduced in Computer Graphics in 1984 by Goral and al. [GTG84]. The original radiosity method is based on a subdivision of the environment into flat polygons. It assumed that all surfaces are perfect Lambertian reflectors and radiosity is constant across the surface of each patch. The computation of form factors is the most time-consuming part of this method. Several papers have proposed various methods to reduce this cost [CGI86], [CCW88], [HSA91].

Another class of radiosity algorithms [BNN98], [FP93], [JP97], [JP98], [JPS99], [NNB97], [Sbe93], [Shi91] tries to reduce the computational cost by using Monte Carlo techniques based on shooting rays. The new method presented in this paper belongs to this class of radiosity algorithms.

The pyramidal hemisphere subdivision method is based on the hemisphere subdivision method [JP97], [JP98]. The hemisphere subdivision method is a

technique that improves a classic Monte Carlo radiosity technique. With the hemisphere subdivision method, a useful image of a scene can be obtained much sooner than with the Monte Carlo radiosity technique. The hemisphere surrounding a surface patch is divided into spherical triangles and interesting directions are searched in each spherical triangle. A direction is interesting if its density (that is the number of surfaces in this direction) is important. Thus, rays are shot to determine interesting directions and each spherical triangle is recursively divided into four spherical triangles. Additional random rays are shot into each spherical triangle. The main drawback of this method is that rays have two different functions: the first one is to compute good directions for shooting rays; the second is to expand the energy of the patch. Thus, some rays do not transport energy while others transport energy proportional to the area of the corresponding spherical triangle. Another drawback of the method is its lack of accuracy, because the density of a region of the scene is an average value computed from densities according to a small number of directions.

Pyramidal hemisphere subdivision, initially presented in [JPS99], is an alternative to the hemisphere subdivision. As, in the hemisphere subdivision method, rays have two different functions, it would be interesting not to use rays at all in the good direction determining process. So, triangular pyramids are used to compute densities in a global manner.

In section 2, the general principle of algorithms using Monte Carlo techniques to compute radiosity will be presented. In section 3, we will give a complete presentation of the pyramidal hemisphere subdivision method and will present some improvements of the initial version, while results of this new method will be given in section 5. Implementation details will be explained in section 4. We will conclude in section 6.

2. MONTE CARLO RADIOSITY

The Monte Carlo based algorithms work according to the following principle: given a patch i and a set of N_i points uniformly distributed over the surface of the patch, a direction is generated for each point of the patch, according to the cosine of the angle between this direction and the patch normal. For a patch j , the number N_{ij} of lines which have the nearest intersection point with the surfaces of the scene on the patch j , can be calculated. So, the ratio N_{ij}/N_i is an approximation of the form factor F_{ij} between patches i and j .

The pseudo-code program in Figure 1 describes a conventional Monte Carlo method [Shi91], [FP93], where Φ_i^{unsent} is the accumulated energy not yet propagated of the patch i , Φ_i^{emit} is the exitance and Φ_i is the total energy leaving the patch i . In Monte Carlo radiosity methods the energy of a patch is divided in r rays that carry the same energy $_{ray}$. The number of rays sent from a patch i is directly proportional to Φ_i^{unsent} .

In Figure 2, we describe the shooting rays algorithm for a conventional Monte Carlo method, where r_j is the reflectivity of the patch j .

The radiosity computing process is terminated when all the energy of the scene has been distributed or when a useful image has been obtained.

A problem with the Monte Carlo radiosity is the noise due to its inability to fit into different kinds of scenes in order to increase the convergence. So, the process of distributing the energy of a patch is independent of the position of the other patches in the scene. To improve Monte Carlo radiosity, we present a new method that takes into account the properties of the objects (position, reflectivity...) to

distribute energy more precisely, in order to obtain a useful image more quickly.

```

for all patches  $i$  do
   $\Phi_i^{unsent} = \Phi_i - \Phi_i^{emit}$ 
while no convergence do
{
  for all patches  $i$  do
  {
    Compute  $r_i$ , the number of
      rays to shoot from patch  $i$ 
    Send the rays in order to
      propagate energy
     $\Phi_i^{unsent} = 0$ 
  }
}

```

Figure 1: Pseudo-code describing a conventional Monte Carlo radiosity method

```

 $_{ray} = \Phi_i^{unsent} / r_i$ 
for the  $r_i$  rays do
{
  Choose a random point,
    uniformly distributed on  $i$ 
  Choose a random direction  $(\theta, \phi)$ ,
    generated according to the
    cosine law
  Find the nearest intersected
    patch  $j$ 
   $_{j}^{unsent} = _{j}^{unsent} + _{j} * \phi_{ray}$ 
}

```

Figure 2: Pseudo-code describing the shooting rays process in a conventional Monte Carlo method

3. REGULAR HEMISPHERE SUBDIVISION BY USING A PYRAMID

3.1 Previous work

The hemisphere subdivision method is an adaptive technique permitting to improve the progressive refinement Monte Carlo radiosity [FP93] by optimizing the choice of shooting rays [JP97], [JP98].

More precisely, the purpose of this method is to avoid tracing useless rays for visibility computation, in order to obtain a good image sooner, while sending more rays towards complex regions in order to obtain a more precise image. To do this, the algorithm works in the following manner: the hemisphere surrounding a surface patch is divided into four spherical triangles and interesting directions are searched into each spherical triangle. A direction is interesting if its density (that is the

number of surfaces in this direction) is significant. Thus, rays are shot to determine interesting directions and each spherical triangle is recursively divided into four new spherical triangles.

Two techniques to distribute the energy have been implemented. The first method propagates the energy with the rays used to estimate the densities. The second method uses two steps, a preprocessing phase and a distribution phase.

3.2 Pyramidal hemisphere subdivision

As for the hemisphere subdivision method, the main idea of the new method presented in this paper is to determine regions of the scene where the distribution of energy must be done more precisely (see Figure 3) because the scene is more complex in these regions.

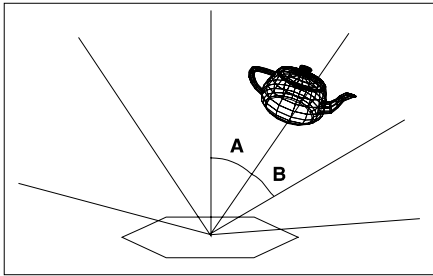


Figure 3: *Distribution of energy must be more precise in regions A and B*

To identify all regions of the scene from a patch, we construct a hemisphere on the patch surrounding its center. The hemisphere is subdivided into spherical triangles.

Then, if the regions of the scene have to be subdivided more precisely, each spherical triangle is subdivided into smaller spherical triangles. The subdivision criterion is the density of region, but a new definition of the density of a region is used in this new method, allowing more accurate estimation.

To subdivide a hemisphere, various techniques can be used. A first technique is based on hemisphere subdivision in spherical triangles with equal areas: the hemisphere is divided into four equal spherical triangles and then, each spherical triangle is subdivided into three new smaller, equal sized, spherical triangles by joining the center of the current spherical triangle with its three vertices. The main drawback of this technique is that the subdivision process cannot be recursively repeated because the areas of the spherical triangles should not be equal (see Figure 4a).

Another implemented technique is to subdivide each spherical triangle into four new spherical

triangles. This is obtained by joining the middle points of each edge of the current spherical triangle (see Figure 4b). This method has the advantage to be recursive, but the areas of the new spherical triangles are not all equal (see Figure 5)

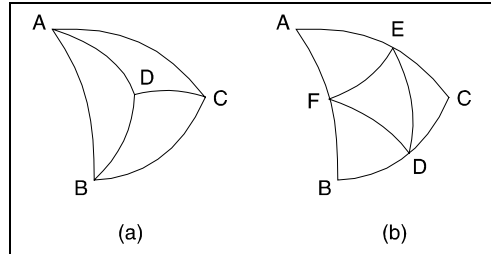


Figure 4: *The initial spherical triangle ABC is divided into three or four new spherical triangles*

At the end of the subdivision process, N_{tr} triangular pyramids are obtained, sampling the scene into N_{tr} regions. Each pyramid is defined by the center of the patch and three planes. Each plane is defined by two vectors. These vectors have as origin the center of the patch and as direction a vertex of a spherical triangle (see Figure 6).

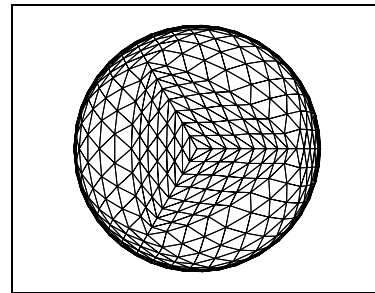


Figure 5: *The areas of spherical triangles are not equal*

For each pyramid, its density is calculated. The density of a pyramid determines the fraction of rays to send in the pyramid in comparison with the number of rays sent by the patch in the hemisphere. In conventional Monte Carlo methods, the number of rays to shoot in a region is proportional to the cosine of the region with the normal of the patch. In our method, the density of a region depends on the cosine of the region and on the number of objects contained in this region.

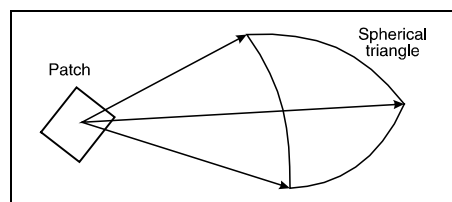


Figure 6 : *Triangular pyramid associated with a spherical triangle*

The general strategy to calculate the density is the following. The density of a pyramid is function of two criteria. The first criterion is the ratio of the number of patches inside the region defined by the plane of the patch and its normal, and of the number of patches contained in the pyramid. The second criterion is the fraction of energy contained in the pyramid. So, for a hemisphere surrounding a patch i divided in N_{tr} spherical triangles, we calculate the density of a pyramid p using the following formula:

$$density_{ip} = hf(f_p, n_{ip})$$

$$\sum_{j=1}^{N_{tr}} density_{ij} = 1$$

where $density_{ip}$ is the density in the pyramid p for the patch i , f_p is the fraction of energy of the patch i propagated in the pyramid p , n_{ip} is the number of patches contained by p and hf is a heuristic function.

The fraction of energy propagated in a pyramid is independent of the patches.

$$f_p = f_i$$

Thus, we can calculate and store this fraction f_p for all the pyramids of the hemisphere.

The heuristic function depends on the fraction of energy f_p and on the number of patches n_{ip} , contained into a pyramid. We can define a general form of this function:

$$hf(f_p, n_{ip}) = \frac{\alpha_1 f_p + \alpha_2 g(n_{ip})}{\alpha_1 + \alpha_2}$$

where $g(n)$ is the ratio of the patches contained in a pyramid:

$$g(n_{ip}) = \frac{n_{ip}}{\sum_{j=1}^{N_{tr}} n_{ij}}$$

In this paper α_1 and α_2 are equal, so the heuristic function used becomes:

$$hf(f_p, n_{ip}) = \frac{f_p + g(n_{ip})}{2}$$

The general algorithm of this method could be decomposed in two parts. The first part is a preprocessing phase where we estimate the density of each pyramid for all the patches. The second part is a conventional Monte Carlo method (see Figure 1) with a new method to shoot rays described in Figure 7 where r_{ip} is the number of rays to shoot in the pyramid p from the patch i . This number is proportional to the density of the pyramid. The number of rays shot from patch i is approximately the same as in the Monte Carlo radiosity method, but their distribution is different.

```

For all the pyramids p do
{
  rip = densityip * ri
  for the rip rays do
  {
    Choose a random point,
    uniformly distributed on i
    Choose a random point on the
    spherical triangle p
    Find the nearest intersected
    patch j

    j1 j2 j3 * φray
    unsendj unsendj unsendj * φray
  }
}

```

Figure 7: *Pseudo-code describing the shooting rays process in the Pyramidal Hemisphere Subdivision method*

The implementation of the spherical triangle sampling is based on Arvo's method [Arv95]. According to this method, random points uniformly distributed on the spherical triangle surface are chosen. To keep an accurate distribution of the energy, we must calculate the energy sent by a ray. This energy is proportional to the cosine of the angle between the direction of the ray and the normal of the patch.

The distributed energy tr in a pyramid defined by a spherical triangle is:

$$tr = \int_{\Omega_r} \cos d \frac{1}{\pi}$$

Now, selecting uniformly a point on the triangle to obtain a direction is the same as solving by Monte Carlo integration the above integral with the following uniform pdf:

$$pdf = \frac{1}{\int_{\Omega_r} d} = \frac{1}{\Delta_r}$$

This means that, if we select points, the value of tr is approximated by:

$$tr = \sum_{j=1}^{N_{tr}} \frac{i_{tr} \cos j}{\pi r_{tr}}$$

This can be done by choosing the energy sent by a ray in the spherical triangle as follows:

$$ray = \frac{i_{tr} \cos}{\pi r_{tr}}$$

where r_{tr} is the number of rays shot in the pyramid, and i_{tr} is the angle between the direction of the ray and the normal of the patch.

If the used unit hemisphere is subdivided in N_{rr} spherical triangles with equal area, the above two formulae become:

$$r_{rr} = \sum_{j=1}^{N_{rr}} \frac{2 \cos \theta_j}{N_{rr} r_{rr}}$$

$$r_{ray} = \frac{2 \cos \theta}{N_{rr} r_{rr}}$$

3.3 Optimization of the pyramidal hemisphere subdivision method

In the Pyramidal Hemisphere Subdivision Method, we calculate a density for each pyramid. To calculate this density we must estimate the number of patches which can receive some energy. In the Section 3.2, this number is the number of patches contained in a pyramidal region.

The example of the Figure 8 shows that the number of patches contained in a region can be a very rough estimation of the number of visible patches, because a patch occludes a region.

A better approximation of this number can not be obtained in the pre-processing phase because this computation would be very time-consuming.

The idea of this optimisation is to use the rays shot from the patch (see Figure 7) to refine the estimation of the number of patches which have received at least one ray. For each pyramid, we calculate the ratio α_{ip} of patches intersected by at least one ray.

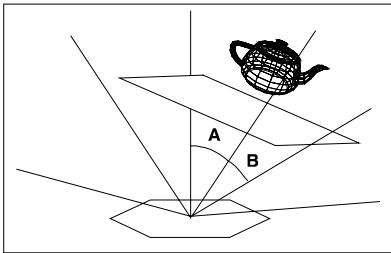


Figure 8: The regions A and B are occluded by a patch

In order to calculate the $density_{ip}$, we change the function g which becomes:

$$g(n_{ip}) = \frac{\alpha_{ip} n_{ip}}{\sum_{j=1}^{N_{rr}} \alpha_{ij} n_{ij}}$$

Now the value of the $g(n_{ip})$ changes during the resolution process. This optimisation improves the efficiency of the Pyramidal Hemisphere Subdivision method.

3.4 Relation with importance methods

Importance methods work by assigning an initial importance V to a whole region in object space or in image space (i.e. the visible part of the scene). The importance I is then computed (usually on the fly or with a progressive approach [NNB96], [PM93], [Sbe97]) as the solution of the dual to radiosity system of equations:

$$I_i = \sum_j R_j F_{ji} I_j + V_i$$

The form factors are then weighted with the importances to obtain transition probabilities biased towards patches that are important for the region of interest. This procedure however has the drawback of having to pre-compute the form factors [PM93].

A variation is *directional* importance [NNB96]. In it we are interested in which directions are more important for the illumination of a region, thus the importance of a patch is substituted by the importance of a hemispherical region. That is, a hemispherical region is important if visible patches within this region are important for the region of interest. The hemispherical regions that have been used so far are obtained with parallel and meridian subdivision. Now, our method bears a resemblance with directional importance, but there are important differences. First is that our region of interest is not a connected region. This is, small patches or high density of objects can be distributed all along the scene, thus making impractical a directional importance approach: all hemispherical regions would have almost the same importance!. Second is that our method is a greedy one: we only mind for our regions being *directly* important to our unconnected regions of interest, i.e. small patches or high density of objects. We do not mind at all if a hemispherical region is *indirectly* important, this is, if it sees a patch that in its way is important for the region of interest. This greedy approach permits us to really concentrate work where it is needed.

4. IMPLEMENTATION

The Pyramidal Hemisphere Subdivision Method has been implemented in a Digital Alpha 500au. In the implemented version, the patches of the scene are grouped into clusters [Chr95]. If a cluster is, totally or partially, contained into a pyramid, the clustered patches are tested for belonging to the pyramid. A patch belongs to a pyramid if, at least, one vertex of the patch is contained into the pyramid. The patches with null reflectivity are not taken into account.

The initial hemisphere is currently subdivided into four spherical triangles but this number can be modified by the user. The two spherical triangle subdivision techniques viewed in section 3.2 have

been implemented and the user can choose the technique he (she) prefers.

5. RESULTS AND DISCUSSION

The Pyramidal Hemisphere Subdivision (PHS) method has been compared with the Progressive Monte Carlo Radiosity (PMCR) method [FP93].

The first test scene contained 13073 patches and 10950 vertices. Experimental tests show that the Pyramidal method improves Feda's method by obtaining a useful image faster than the latter one. In Figure 11 one can see images obtained by the two methods after five processing steps and 176158 rays sent in Feda's method while 179281 rays were sent in the Pyramidal Hemisphere Subdivision method.

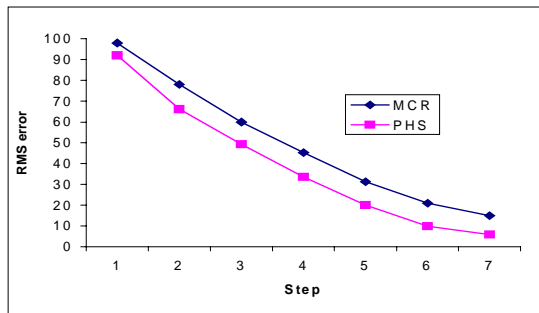


Figure 9: Evolution of the Global RMS Error for the two methods

The evolution of global RMS error, corresponding to the whole scene, is described by the graph of Figure 9, while Figure 10 shows the evolution of a local RMS error for a part of the scene formed by the table and the nearest chair (see Figure 12). In these graphs MCR means "Monte Carlo Radiosity" and PHS "Pyramidal Hemisphere Subdivision".

The time is practically the same for the two methods, the additional pre-processing time being negligible for the scenes used. The additional memory required by PHS method is approximately of 1 Mb for the test scene.

The second test scene contained 1450 patches and 1186 vertices. Figure 13 shows three images. The first one is the reference scene while the others represent images of a part of the scene, rendered using the PMCR method and the PHS method respectively. The number of rays shot is approximately 100000 rays for the two methods.

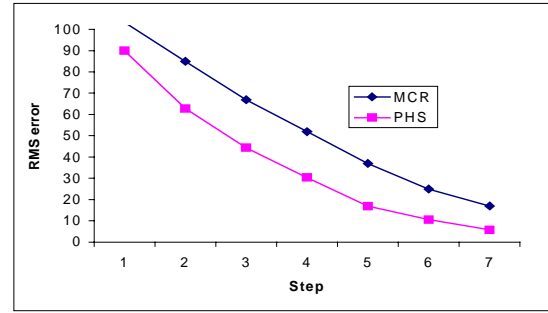


Figure 10: Evolution of the Local RMS Error for the two methods

We expected to render the fine details much sooner and with lower cost, but we surprisingly also obtained a considerable increase in the overall efficiency. This we interpret as a consequence of the more clever ray distribution (and hence, of the energy).

6. CONCLUSION AND FUTURE WORK

A formal description of the pyramidal hemisphere subdivision technique, has been presented in this paper. This technique permits to obtain a useful image of the scene much sooner than with the progressive refinement Monte Carlo radiosity algorithm. Moreover, the pyramidal subdivision technique permits a more accurate processing of complex regions of the scene.

An improvement of the Pyramidal Hemisphere radiosity method, allowing more accurate calculation of the density by using a dynamic estimation of the number of visible patches in a region, has also been described.

An adaptive version of the subdivision technique has been studied and implemented.

We are actually working on the implementation of this method in the Hierarchical Monte Carlo Radiosity [BNN98].

7. ACKNOWLEDGEMENTS

We would like to thank the authors of RenderPark, a platform to integrate radiosity algorithms, to have authorized us to use their scenes in order to compare our methods with other methods. The third author has been funded in part with a CICYT grant number TIC98-0586-c03-02.

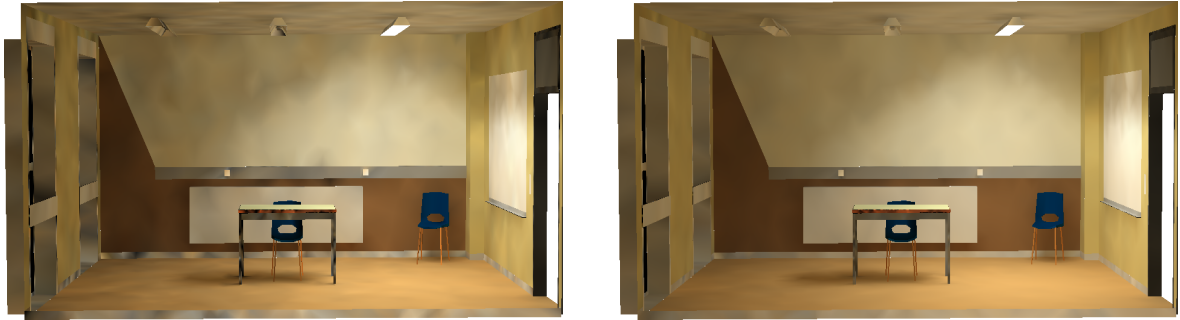


Figure 11: A scene rendered with Monte Carlo radiosity method (left), and with the new method (right)

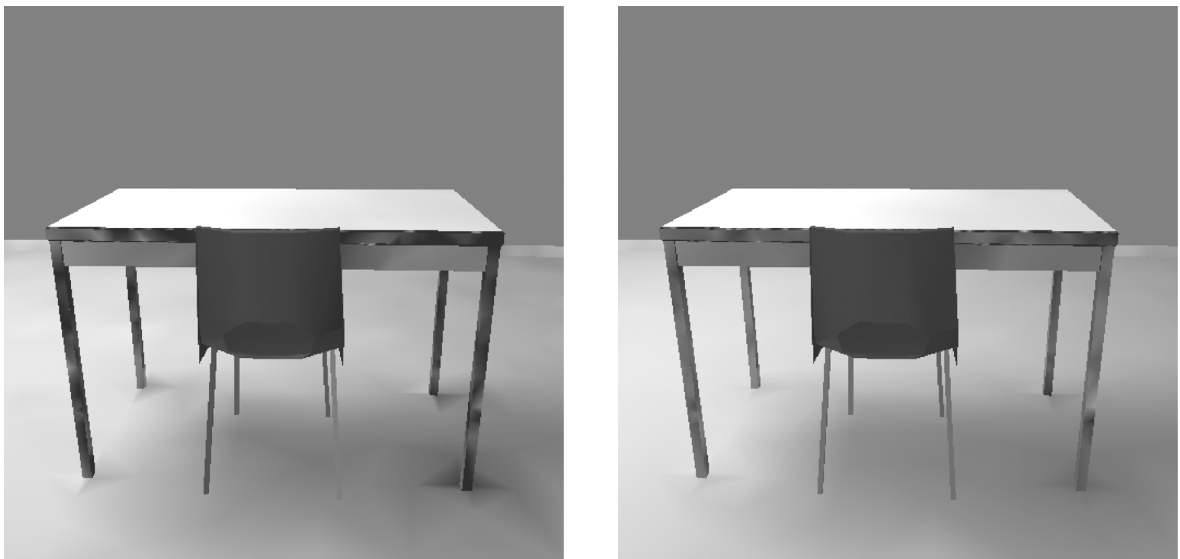


Figure 12: Details of the scene of Figure 11 rendered with Monte Carlo radiosity (left), and with the Pyramidal Hemisphere Subdivision method (right)



Figure 13: Another scene and details rendered with Monte Carlo radiosity (middle), and with the Pyramidal Hemisphere Subdivision method (right)

REFERENCES

- [Arv95] Arvo J., Stratified sampling triangles. In *Computers Graphics*, volume 29, pages 431-438, August 1995.
- [BNN98] Bekaert P., L. Neumann, A. Neumann, M Sbert, Y.D. Willems, Hierarchical Monte Carlo Radiosity. In *Eurographics Workshop on Rendering*, pages 259-268, June 1998.
- [Chr95] Christensen P., Hierarchical Techniques for Glossy Global Illumination, PhD Thesis, University of Washington, page 116, 1995.
- [CCW88] Cohen M., S.E. Chen, J.R. Wallace, D.P. Greenberg, A Progressive refinement approach to fast radiosity image generation. In *Computer Graphics(ACM SIGGRAPH'88 Proceedings)*, volume 22, pages 75-84, August 1988.
- [CGI86] Cohen M., D.P. Greenberg, D.S. Immel, P.J. Brock, An efficient radiosity approach for realistic image synthesis. *IEEE Computer Graphics and Applications*, 6(3):26-35, March 1986.
- [FP93] Feda M., W. Purgathofer, Progressive Ray Refinement for Monte Carlo Radiosity. In *Fourth Eurographics Workshop on Rendering*, pages 15-25, June 1993.
- [GTG84] Goral G.M., K.E. Torrance, D.P. Greenberg, B. Bataille, Modeling the Interaction of light Between Diffuse surfaces. In *Computer Graphics (ACM SIGGRAPH'84 Proceedings)*, volume 18, pages 212-222, July 1984.
- [HSA91] Hanrahan P., D. Salzman, L. Aupperle, A rapid hierarchical radiosity algorithm. In *Computer Graphics (SIGGRAPH'91 Proceedings)*, volume 25, pages 197-206, July 1991.
- [JP97] Jolivet V., D. Plemenos, The Hemisphere Subdivision Method for Monte Carlo Radiosity. In *Graphicon'97*, Moscow, 21-24 May 1997.
- [JP98] Jolivet V., D. Plemenos, A New Hemisphere Subdivision Method for Monte Carlo Radiosity. In *Graphicon'98*, Moscow, 5-12 September 1998.
- [JPS99] Jolivet V., D. Plemenos, M. Sbert, A Pyramidal Hemisphere Subdivision Method for Monte Carlo Radiosity. In *Proc. Eurographics '99, Short papers*, Milano, September 1999.
- [NNB97] Neumann L., A. Neumann, P. Bekaert, Radiosity with well-distributed ray sets. In *Computer Graphics Forum*, 16(3):C261-269, 1997.
- [NNB96] Neumann A., L. Neumann, P. Bekaert, Y. Willems, W Purgathofer,, Importance-Driven Stochastic Ray Radiosity. In Proceedings of the Seventh Eurographics Workshop on Rendering. Appeared in *Rendering Techniques '96*, pages. 111-122, Springer, Wien, 1996.
- [PM93] Pattanaik S., and S.Mudur, Efficient potential equation solutions for global illumination computation. In *Computers & Graphics*,17(4):387-396, 1993.
- [PP96] Plemenos D., X. Pueyo, Heuristic Sampling Techniques for Shooting Rays in radiosity algorithms. In *3IA'96 Conference*, Limoges (France), 3-4 April 1996.
- [Sbe93] Sbert M., An integral geometry based method for fast form-factor computation. In *Computer Graphics Forum*, 16(3):C409-420, 1993.
- [Sbe97] Sbert M., Optimal source selection in shooting random walk radiosity. In *Computer Graphics Forum (Proceedings of Eurographics'97)*, 16(3), pages C301-C308, Budapest, September 1997.
- [Shi91] Shirley P., Radiosity via raytracing. In James Arvo, editor, *Graphics Gem II*, pages 306-310. Academic Press, San Diego, 1991.