A New Hemisphere Subdivision Technique For Computing Radiosity

Vincent JOLIVET, Dimitri PLEMENOS

jolivet@unilim.fr, plemenos@unilim.fr

MSI laboratory, University of Limoges, France.

ABSTRACT

The hemisphere subdivision technique is an importance-driven technique using a triangular adaptive subdivision of a hemisphere in order to guide the rays shot from a patch [17]. The purpose of this technique is to take advantage of the knowledge of an environment in order to optimise the space distribution of rays shot in radiosity algorithms. In this paper, we present a new approach of the hemisphere subdivision technique, allowing both reduction of time and memory costs of the technique and improved results. Like the initial hemisphere subdivision technique, this approach, uses a heuristic function which estimates visibility complexity in a given direction from a patch. The knowledge of this complexity is used as a guide for subdividing a hemisphere associated with the patch, in a number of equilateral spherical triangles of various sizes. Then, during a number of steps chosen by the user, new rays are sent in specific directions derived from the hemisphere subdivision.

KEYWORDS

Radiosity, visibility, heuristic search, progressive refinement, Monte-Carlo radiosity, form factors.

1 INTRODUCTION

Radiosity has been introduced in computer graphics in 1984 by Goral et al. [5] borrowed from thermal engineering for realistic image synthesis of scenes with diffuse inter-reflections. For radiosity computation, the form factor between each pair of elements of the scene (patches) must be obtained. This implies computing visibility between each pair of patches and is responsible for about 80% of the cost of the algorithm. The results of research in this area have enhanced the applicability of radiosity. In this direction we may underline adaptive subdivision [1], progressive refinement [2] and hierarchical radiosity [6]. Nevertheless, visibility computation remains the important parameter in the cost of radiosity algorithms.

To the authors' knowledge, up to now it has not been proposed any strategy which takes really advantage of the knowledge of the scene in order to either reduce the number of rays or optimize the order of tracing them for the computation of visibility (form factors), with the exception of the papers of Plemenos and Pueyo [10] and Plemenos and Jolivet [17] where heuristic methods were proposed, based on the techniques proposed for automatically finding *good view directions* [7, 3, 9]. Some authors have understood the importance of this problem but the proposed solutions are not entirely satisfactory. Thus, two papers have introduced the concept of visibility complex for efficient computation of 2D and 3D form factors [13, 18]. Another paper [19] introduces the notion of importance-driven stochastic ray radiosity, an improvement of the stochastic ray radiosity. The purpose of this paper is to improve the hemisphere subdivision technique published in [17] and applied to the the family of radiosity computation methods which uses Monte-Carlo techniques. Indeed, although the hemisphere subdivision method has several advantages, its main drawback is a very important cost in memory occupation. The improvements presented in this paper permit to obtain a useful image sooner than with classical methods, with an acceptable memory cost.

In section 3 we will describe some well known radiosity computing techniques, after recalling in section 2 its basic concepts. Section 4 describes the principle of the hemisphere subdivision method as it has been presented in [17]. In Section 5 we describe the new approach of the hemisphere subdivision method, improving its memory and time complexity. Some results of the new approach are shown in section 6.

2 RADIOSITY

The radiosity algorithm is a method for evaluating light intensity at discrete points of ideal diffuse surfaces in a closed environment [5]. This environment is composed of polygons which are split into patches. The relationship between the radiosity of a given patch and the remainder is given by the following formula:

$$B_i = E_i + P_i \sum_{j=1}^{n} F_{ij} B_j$$

where :

Bi = radiosity in surface i
Ei = emissivity of surface i
Pi = reflectivity of surface i
Fij = form factor between surface i and j
n = number of surfaces in the environment

A form factor is given by the geometric relationship

between two patches and represents the ratio of energy leaving one and arriving to the other [5].
The form factor between finite patches is defined by:

$$F_{A_i A_j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos\alpha_i \cos\alpha_j \, HID}{\pi r^2} \, dA_j \, dA_i$$

where function HID represents the visibilitty between the patches *Ai* and *Aj*.
The progressive approach [2] avoids the enormous cost of radiosity by computing form factors on the fly. Now, instead of keeping the whole system of equations to complete the energy exchange of the environment, a useful solution is computed. This is obtained by only shooting the energy of a reduced number of patches, the ones which contribute the most to the illumination of the environment.

3 PROGRESSIVE REFINEMENT AND MONTE-CARLO BASED TECHNIQUES

As stated in the previous section, radiosity is computed in environments composed of surfaces which are divided into patches. The purpose of radiosity improvement techniques is to obtain as fast as possible a useful image. Most of radiosity computation techniques are progressive techniques, where the radiosity of a scene is computed and refined in several steps.

3.1 Progressive Radiosity

The energy of each patch is distributed, starting from patches which contribute the most to the illumination of the environment [2]. The form factors are explicitly computed between the emitting patch and each one of the other patches of the environment.
An image of the scene is displayed each time a patch has been processed. So, the displayed image is progressively improved.
The radiosity computing process is achieved when all of the scene's energy has been distributed. It is also possible to stop the radiosity process if a useful image is obtained.

3.2 Monte-Carlo Radiosity

Starting from the patch A_i which possesses the maximum of energy ($A_i.B_i$), a great number of rays is sent from randomly chosen points of each patch towards directions chosen according to a cosine distribution [11]. Each ray distributes the same amount of energy. The rays that are shot are supposed to distribute the whole energy of the patch.
Form factors are not explicitly computed. The fraction of the energy received by a patch is proportional to the number of received rays and to the energy diffused by each received ray.
An image of the scene is displayed each time the whole energy of a patch has been distributed. So, the displayed image is progressively improved.
The radiosity computing process is achieved when all of the scene's energy has been distributed or when a useful image has been obtained.

3.3 Monte-Carlo Progressive Radiosity

An additional loop is added to the pure Monte-Carlo algorithm. At each step, the energy of a reduced number of additional rays is shot from each patch of the scene [4]. An image is displayed only at the end of a step, that is when the whole energy of a reduced number of rays has been totally distributed in the scene.
If the total number of rays shot until the current step (included) is k times greater than the number of rays shot until the previous step, the energy received by each patch until the previous step and the energy distributed by the current patch are divided by k and the remaining energy is then shot by the additional rays sent in the current step.
The whole process can be stopped if the produced image is considered useful.

4 THE HEMISPHERE SUBDIVISION TECHNIQUE

The technique presented in this section is an adaptive technique permitting to improve the progressive refinement Monte-Carlo radiosity by optimising the choice of shooting rays. This approach, which was initially presented in [17], divides a spherical triangle using a heuristic based on the density of the scene in a given direction. The target of this technique is to reduce, at least in a first step, the number of rays shot from an emitter in order to obtain a useful image.
In the progressive refinement approach proposed in [4] for Monte Carlo radiosity, the algorithm starts distributing the energy using a reduced number of rays. Afterwards the number of rays is increased in order to refine the solution.
So it would be interesting to find "intelligent" algorithms which avoid tracing useless rays for visibility computation, in order to obtain a good image sooner.
The goal is to reduce the number of rays by performing a *good* choice of them. Thus, it is possible that in some directions a reduced number of rays is enough while in other directions more rays are necessary. The chosen strategy is to send more rays towards the directions where more surfaces will be found. This strategy is justified by the fact that *the probability to meet a surface is greater for a ray sent to a direction where there are many surfaces than for a ray sent to a direction where the number of surfaces is small.* Thus, let us consider the hemisphere surrounding the emitting patch (figure 1).

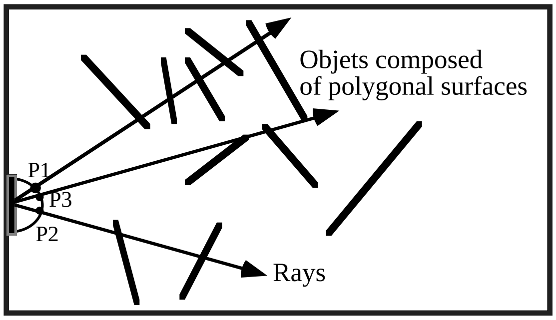

Figure 1 : Casting selected rays.

The following method is used to select new rays to shoot: if the number of surfaces (patches) intersected by a ray through point P1 in the hemisphere is ns(P1) and the number of surfaces intersected in the direction defined by point P2 is ns(P2); then a new ray will be traced into the direction defined by P3 where P3 verifies:

$$\frac{P_3P_1}{P_2P_3} = \frac{ns(P_2)}{ns(P_1)} \qquad (1)$$

This process is recursively repeated.

The above strategy may be generalised to 3D. Now the hemisphere is divided into four spherical triangles. Each spherical triangle is processed in the same manner: a ray is traced for each vertex from the centre of the patch in order to determine the number of surfaces it intersects (ns(A) for vertex A). Then, on each edge, a new point is computed, using the above formula (1), which determines the direction of a new ray to shoot. These three new points, together with the initial vertices of the spherical triangle, define 4 new spherical triangles which are processed in the same manner (figure 2).

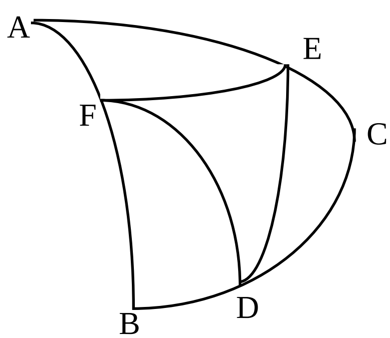

Figure 2 : Computation of 4 new spherical triangles.

In practice, this naive hemisphere subdivision method is very expensive in memory occupation because it is difficult to store all the spherical triangles at each iteration. So, the implemented algorithm is slightly different and works as follows:

1. For each patch, rays are traced from a small number of points of the patch, chosen by a fictive regular subdivision of the patch, using the hemisphere subdivision method described in section 4. A hemisphere is computed for each point (figure 3).

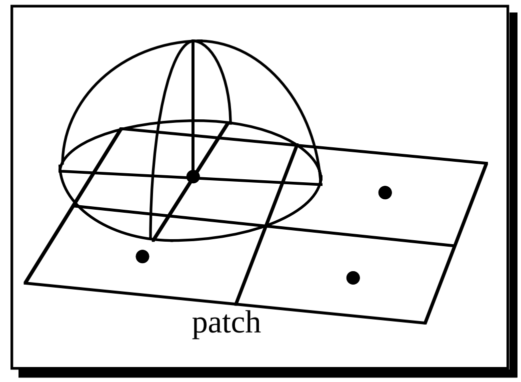

Figure 3 : fictive patch subdivision and hemisphere computed for each sub-patch centre

After each spherical triangle refinement, the same number of additional rays is traced randomly to each spherical triangle. This hemisphere subdivision is pursued during a number **n** of steps given by the user. The obtained spherical triangles are stored at each step, in order to be subdivided in the following steps.

2. During another number **m** of steps, also given by the user, the same subdivision process is applied but now the spherical triangles are not stored. They are approximately estimated at each refinement step. The approximate estimation process works as follows, where ens (A) means "estimated number of surfaces encountered in direction A" :

(i) Initialise estimated numbers of surfaces for directions A, B and C :
ens(A) = ns(A),ens(B) = ns(B), ens(C) = ns(C).

(ii) At any subsequent refinement step, compute, from the current spherical triangle ABC, estimated values for new directions D, E and F. The following formulas give the way in which the estimated position of the new point F, and the estimated number of visible surfaces from this point are computed :

$$\frac{AF}{FB} = \frac{ens(B)}{ens(A)}$$

$$ens(F) = \frac{ens(A)\,BF + ens(B)\,AF}{AB}$$

3. Traditional Monte-Carlo progressive refinement is applied until a useful image is obtained.
Rays shot towards a spherical triangle transport an amount of energy proportional to the area of this triangle.

5 A NEW APPROACH OF HEMISPHERE SUBDIVISION

The hemisphere subdivision technique presented above is very interesting because it allows intelligent selection of rays shot from a patch during the radiosity computation process. All the tests we have made with the Feda's and Purgathofer's algorithm confirm the improvements due to this technique.
However, we have seen that the brute hemisphere subdivision technique is a memory consuming technique and so, this technique cannot be applied to scenes divided into an important number of patches. For this reason, in the version of the hemisphere subdivision technique really implemented, the subdivision of spherical triangles becomes fictive after a small number of initial subdivisions and the computation of the scene density at vertices of fictive spherical triangles is an approximate one.
The purpose of the new approach presented in this section is to make the hemisphere subdivision technique as accurate as possible, with a memory cost allowing to process complex scenes.

The main principles which guide this approach are the following :

- Separation of the role of rays. The rays shot to determine regions do not distribute energy.

- Separation of the subdivision process from the radiosity computing process.

- Increase of the accuracy of the technique by avoiding fictive triangle subdivisions and approximate computations of scene densities.

- Decrease of the memory cost by reducing the maximum number of subdivisions of a spherical triangle.

After having presented the main purposes of the new hemisphere subdivision method, let us describe the method more precisely.

a. Preprocessing.
A hemisphere, divided into 4 initial spherical triangles, is associated with each patch of the scene (figure 4).

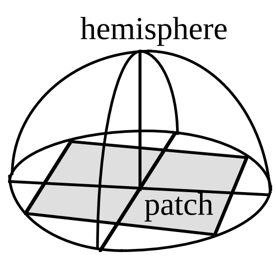

Figure 4 : Initial subdivision of a hemisphere

Each hemisphere is processed in the following manner :

1. From each centre of patch, three rays are shot to each spherical triangle, one ray per vertex, in order to determine the scene's density in the direction defined by the centre of the patch and the vertex (figure 5).

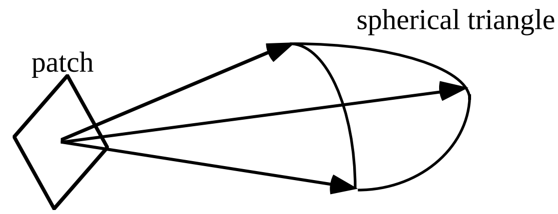

Figure 5 : One ray is sent to each vertex of the triangle

2. If the average value of densities for a spherical triangle is higher than a threshold value, the spherical triangle is divided into four new equal spherical triangles. In figure 6 a spherical triangle is divided into four new spherical triangles while in figure 7 adaptive subdivision of a spherical triangle is shown at the end of the preprocessing phase.

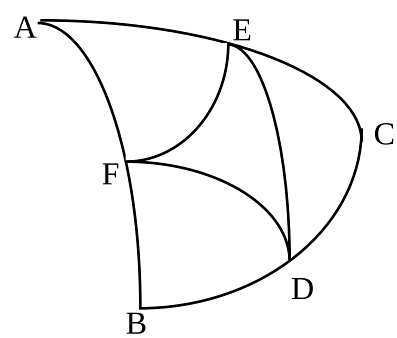

Figure 6 : The initial spherical triangle ABC is divided into 4 equal new spherical triangles

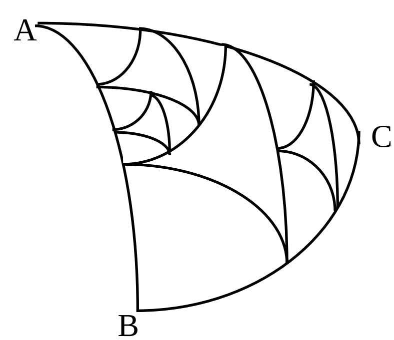

Figure 7 : selective subdivision of a spherical triangle.

Actually, the maximum subdivision depth used for the preprocessing phase is equal to 2 but it can be changed interactively. So, each initial spherical triangle of each hemisphere can be subdivided into a maximum of 16 spherical triangles.

b. Radiosity computation

This phase works during any number of steps and can be stopped if a useful image is obtained.

1. From the centre of each patch of the scene, a number of additional rays are shot towards the remaining of the scene. A number of rays proportional to the average value of its density are shot in each spherical triangle of the hemisphere, towards directions chosen according to a cosine distribution. The amount of energy distributed by each ray shot to a given spherical triangle is the same, proportional to the area of the spherical triangle and inversely proportional to the number of rays shot in the triangle.

2. The radiosity of each patch computed up until the previous step is combined with the radiosity computed in the current step, in order to compute the radiosity of the patch up until the current step.

The general algorithm of the new hemisphere subdivision technique for computing radiosity is briefly presented here:

```
Procedure NewHemisphereRadiosity (Scene)
    Preprocessing (Scene)
    Compute Radiosity (Scene)
end
```

The two procedures *Preprocessing* and *Compute Radiosity* could be written as follows:

```
Procedure Preprocessing (Scene)
    for each patch of Scene do
      for each spherical triangle of
        Hemisphere (patch) do
        if density (spherical triangle) >ThresholdValue
          and SubdivisionLevel < MaxLevel then
          Divide (Spherical Triangle)
      end
    end
end
```

In the procedure *Compute Radiosity*, the meaning of the used variables is the following:

N_{tr}, N_{hemi}: Number of rays to shoot in the spherical triangle and in the hemisphere.

A_{tr}, A_{hemi}: Area of the spherical triangle and ot the hemisphere.

D_{tr}, $D_{average}$: Density of the spherical triangle and average density of the hemisphere.

Φ_{tr}, Φ_{patch}: Energy to shoot in the spherical triangle and total energy to send from the patch.

Φ_{ray}: Energy sent in the spherical triangle by a ray.

Procedure Compute Radiosity (Scene)
  while not Useful (Image) **do**
   for each patch **of** Scene **do**
    for each spherical triangle **of** Hemisphere (patch) **do**

$$N_{tr} := N_{hemi} \frac{A_{tr}}{A_{hemi}} \frac{D_{tr}}{D_{average}}$$

$$\Phi_{tr} := \Phi_{patch} \frac{A_{tr}}{A_{hemi}}$$

$$\Phi_{ray} := \frac{\Phi_{tr}}{N_{tr}}$$

     Choose N_{tr} directions with cosine distribution
     for each direction **do**
      Shoot a Ray (Direction, Φ_{ray})
     end
    end
   end
  end

6 DISCUSSION AND RESULTS

The new hemisphere subdivision method described in section 5 improves the method presented in [17].

The main advantage of this method is to decrease the error committed in regions comporting several details, while the error in regions with few details is not increased significantly. Indeed, the mean square error in regions comporting many details is less than with a classical Monte-Carlo method because more rays are sent towards these regions. In return, although the mean square error in regions with few details is greater than with a classical Monte-Carlo method, this difference is hardly detected by the human eye. So, a useful image is generally obtained very sooner than with the Feda's and Purgathofer's method.

The memory cost of the method is not very important, even for complex scenes. It is increases linearly with the complexity (in number of patches) of the scene.

Thus, if n is the maximum subdivision depth of a hemisphere and p the number of patches of the scene, the maximum number t of spherical triangles we have to process for a hemisphere is given by:

$$t = \frac{4\,(4^{n+1}-1)}{3}$$

For each spherical triangle, the only information we have to store is the average value of its density and this information can be stored in a byte. So, the total memory requirement is:

$$m = p\,t \text{ bytes}$$

or

$$m = \frac{4\,p\,(4^{n+1}-1)}{3} \text{ bytes}$$

Thus, for a complex scene composed of 100 000 patches, a storage capability of 6 400 000 bytes is required.

In figure 7 one can see an image of a simple scene processed with the classical progressive refinement Monte Carlo radiosity method after 3 steps, while in figure 8 we have an image of the same scene processed with the new hemisphere subdivision technique after 3 steps.
The scene is composed of 1002 patches.

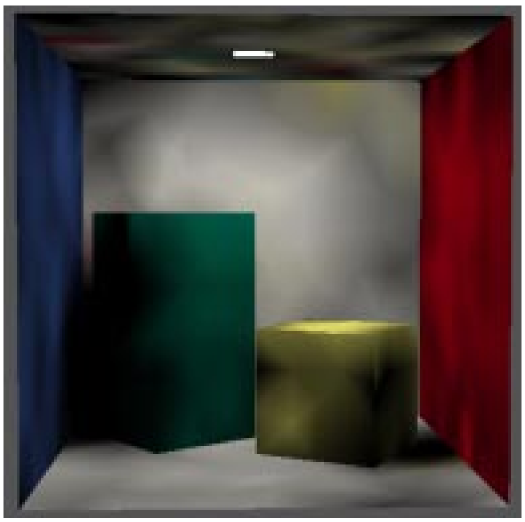

Figure 7: image of a scene after 3 steps with progressive refinement Monte Carlo radiosity

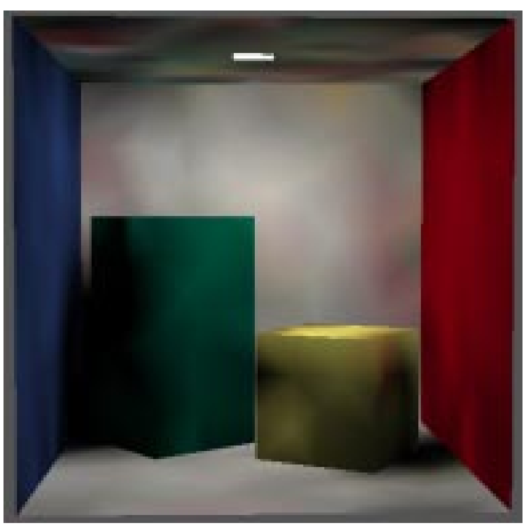

Figure 8: image of a scene after 3 steps with our new hemisphere subdivision method.
The maximum number of spherical triangles of a hemisphere is 84 for a maximum subdivision level equal to 2.

In figure 9 we have an image of a more complex scene processed with the classical progressive refinement Monte Carlo radiosity method after 17 steps, while in figure 10 we have an image of the same scene processed with the new hemisphere subdivision technique after 17 steps.
This scene is composed of 3576 patches and the maximum number of spherical triangles used by the new hemisphere subdivision technique is 84.

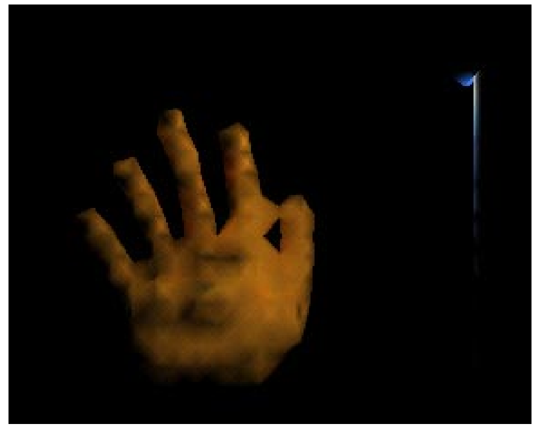

Figure 9: image of a scene after 17 steps with progressive refinement Monte Carlo radiosity

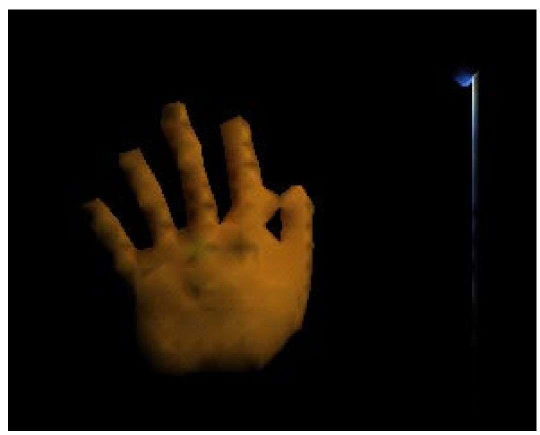

Figure 10: image of a scene after 17 steps with the new hemisphere subdivision method.

7 CONCLUSION

The Hemisphere Subdivision Method is a very interesting importance driven method for computing radiosity. Its main drawback was its cost in memory because of the need to store a big quantity of information about spherical triangles. The new technique of the Hemisphere Subdivision Method we have presented here improves the initial method by significantly reducing the storage requirements.

The main purpose of the technique presented in the paper is to improve the progressive refinement Monte-Carlo radiosity by optimising the radiosity computation in the complex parts of the scene, in order to obtain a useful image of the scene very soon. The results of implementation of this technique show that it permits to obtain a useful image faster than the classical Monte-Carlo progressive refinement technique. However, this technique could possibly fail in some situations where a wall is in front of a large number of surfaces. But, even in such a case, the hemisphere subdivision technique is, at least, as fast as Monte-Carlo progressive radiosity technique.

REFERENCES

[1] COHEN M., GREENBERG D. P., IMMEL D. S., BROCK P. J., An efficient radiosity approach for realistic image synthesis. IEEE Computer Graphics and Applications 6:3 (Mar. 1986), pp. 26-35.

[2] COHEN M., CHEN S. E., WALLACE J. R., GREENBERG D. P., A progressive refinement approach to fast radiosity image generation. Computer Graphics (SIGGRAPH'88 Proc.) 22:4 (Aug. 1988), pp. 75-84.

[3] COLIN C., Automatic computing of scene's "good views". MICAD'90, Paris , February 1990 (in French).

[4] FEDA M., PURGATHOFER W., Progressive Ray Refinement for Monte Carlo Radiosity. Fourth Eurographics Workshop on Rendering, Conf. Proc. (June 1993), pp. 15-25.

[5] GORAL G. M., TORRANCE K. E., GREENBERG D. P., BATTAILE B. Modeling the Interaction of Light Between Diffuse surfaces. Computer Graphics, vol. 18, no 3, (SIGGRAPH'84 Proc.), 1984, pp. 212-222.

[6] HANRAHAN P., SALZMAN D., AUPPERLE L., A rapid hierarchical radiosity algorithm. Computer Graphics (SIGGRAPH'91 Proc.) 25:4 (July 1991), pp. 197-206.

[7] KAMADA T., KAWAI S., A Simple Method for Computing General Position in Displaying Three-Dimensional Objects. Computer Vision, Graphics and Image Processing, 41 (1988).

[8] PATTANAIK S. N., MUDUR S. P., Computation of global illumination by Monte Carlo simulation of the particle model of light. Third Eurographics Workshop on Rendering, Conf. Proc. (May 1992), pp. 71-83.

[9] PLEMENOS D., A contribution to study and development of scene modeling, generation et display techniques - The "MultiFormes" project. Professorial dissertation, Nantes (France), November 1991 (in French).

[10] PLEMENOS D. PUEYO X., Heuristic sampling techniques for shooting rays in radiosity algorithms. 3IA'96 Conference, Limoges (France), 3-4 April 1996.

[11] SHIRLEY P., Radiosity Via Ray Tracing. Graphics Gems II, J. Arvo Ed., Academic Press, San Diego, 1991, pp. 306-310.

[12] WALLACE J. R., ELMQUIST K. A., HAINES E. A., A ray Tracing Algorithm for Progressive Radiosity. Computer Graphics, vol. 23, no 3, (SIGGRAPH'89 Proc.), 1989, pp. 315-324.

[13] ORTI R., PUECH C., Efficient computation of 2D form factors applied to dynamic environments. Annual conference of the AFIG, Marseille, November 1995, pp. 223-230 (in French).

[14] PLEMENOS D., Selective refinement techniques for realistic rendering of three dimensional scenes. Revue Internationale de CFAO et d'Infographie, vol. 1, no 4, 1987 (in French).

[15] JANSEN F. W., VAN WIJK J. J., Previwing techniques in raster graphics, EUROGRAPHICS, September 1994.

[16] COHEN M. F., GREENBERG D. P., The hemi-cube, a radiosity solution for complex environments. SIGGRAPH, 1985.

[17] JOLIVET V., PLEMENOS D., The Hemisphere Subdivision Method for Monte-Carlo Radiosity. GraphiCon'97, Moscow, May 21-24 1997.

[18] DURAND F., DRETTAKIS G., PUECH C., The 3D visibility complex: a new approach to the problems of accurate visibility. Eurographics Rendering Workshop, Porto, 1996.

[19] NEUMANN A., NEUMANN L., BEKAERT P., WILLEMS Y. D., PURGATHOFER W., Importance-driven Stochastic Ray Radiosity. the problems of accurate visibility. Eurographics Rendering Workshop, Saint-Etienne, 1997.

AUTHORS:

Dimitri PLEMENOS, Vincent JOLIVET
University of Limoges
MSI Laboratory
83, rue d'Isle
87000 Limoges
France
Phone: (+33) 555 43 69 74
Fax: (+33) 555 43 69 77
E-mail: plemenos@unilim.fr, jolivet@unilim.fr