A New Mathematical Development for Radiosity Animation with Galerkin
Didier Arquès, Venceslas Biri, Sylvain Michelin

To cite this version:

HAL Id: hal-00681566
https://hal-upec-upem.archives-ouvertes.fr/hal-00681566
Submitted on 21 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A New Mathematical Development for Radiosity Animation with Galerkin

Arquès Didier, Biri Venceslas, Michelin Sylvain
University of Marne-La-Vallée
Institut Gaspard Monge
5, Boulevard Descartes, F-77454 Champs sur Marne, France
{arques,biri,michelin}@univ-mlv.fr

Abstract

Combining animation and global illumination constitutes, at present, a true challenge in computer graphics, especially when light sources move in a complex scene because the entire illumination has to be recomputed. This paper introduces a new algorithm, based on the Galerkin method, which can efficiently manage any moving surface—even light sources—to compute animation sequences. For each new frame of a sequence, we take into account the continuous property of the moves to determine the necessary energy differences between the previous global illumination solution and the new one. Based on a mathematical development of the form factor, this new approach leads to an efficient and simple algorithm, similar to the classical progressive refinement algorithm, and which computes animated sequence about three times faster.

1. Introduction and previous work

The use of realistic global illumination is becoming more and more widespread. But who has never seen, in the actual computer-animated movies, images which seem to be too artificial or shiny, due to inadequate lighting ambiance? This is a consequence of the use of direct illumination algorithms. So there is a need for algorithms that can efficiently combine global illumination and dynamic objects because it really "brings life" into a scene. But it remains one of the most difficult challenges in computer graphics because in a complex scene, dynamic objects (and especially moving light sources) cast new shadows, modify direct illumination and, indirectly, the whole energy relationships between objects.

For static environment, radiosity algorithm constitutes one of the most used means to solve the Kajiya's equation [17], basic equation of any global illumination model. Initially in the traditional radiosity algorithm [8, 13, 22], the radiosity function is projected onto some discrete finite bases. A first animation algorithm, the "back buffer" [2] method, comes directly from this approach and allows to manage any known move with a static camera. Progressive methods [7, 28] lead to new interactive algorithms [5, 12] that propagate light modifications by shooting positive and negative luminous energy. For each move, two steps have to be carried out, one for withdrawing the object and one for adding it, what is unsuitable for moving light source. A more involved data structure has appeared [21] for this progressive refinement radiosity.

Probabilistic methods [18] have also been presented to manage very complex scenes and lead to other animation algorithms [3]. In the same time, the complexity of the function bases used to represent the illumination function in radiosity has been reduced by hierarchical and adaptive methods [4, 15, 19, 20]. Finally other function bases [1, 14, 23, 27, 29, 30] have been found to represent the illumination function but few animation methods based on it have been presented until now. Some animation algorithms [10, 11] have tried to use the benefits offered by adaptive bases, by identifying the links affected by an object displacement.

We present here a new radiosity algorithm, which manages the case of surfaces or light sources in translation. This method takes advantage of the continuity properties of these displacements to compute efficiently all illumination relationships between objects. We obtain a new algorithm similar to classical progressive refinement algorithm with new form factors, which allows to obtain quickly the new global illumination. As any radiosity algorithm, it can also be used as a first pass of a classical rendering system that can add specular and reflection effects.

The next section of this paper briefly returns on the Galerkin algorithm. In the third section, we are presenting the mathematical development and approximations we have used to treat a dynamic scene, while the fourth section shows results obtained by this method.
2. Theoretical Background: the Galerkin Algorithm

Galerkin radiosity [30] is an alternative radiosity formulation that generalizes the classical radiosity approach. This model avoids the last rendering interpolation and allows representing the scene with a restricted number of parametric surfaces. It avoids any discretisation, allows to store every kernel coefficients and separates shadow computation from illumination determination. Unlike the classical radiosity approach, as shown in figure 1, each function used (the radiosity, the exitance and the reflectivity) is projected onto a base of $N'$ orthogonal non-constant functions $\{\Gamma_k(s,t) = 1..N'\}$ defined over each entire surface. For example, Zatz [30] uses a base of Legendre polynomials. Then, the Galerkin method consists in computing, for each surface, the radiosity coefficients associated to each function of the base.

$$ = \Sigma$$

Figure 1. Decomposition on a classical discrete constant function basis and a higher order function basis

More precisely, let us consider a scene defined by $N$ parametric surfaces. For a surface of index $i$ and parameters $(s,t)$, the radiosity equation is :

$$B_i(s,t) = E_i(s,t) + \sum_{j=1}^{N} \int K_{ij}(s,t,u,v) B_j(u,v) \, du \, dv \tag{1}$$

where $E_i$ is the exitance and where the kernel function $K$ is the product of the elementary form factor $F$, the reflectivity $\rho$, the area $A$ and a visibility term $VIS$ :

$$K_{ij}(s,t,u,v) = \rho_i(s,t) F_{i-j}(u,v) VIS_{ij}(s,t,u,v) A_j(s,t) \tag{2}$$

Each function used, respectively the radiosity, the exitance and the reflectivity, is then projected onto a base of $N'$ orthogonal functions $\{\Gamma_k(s,t), k = 1..N'\}$. For instance, radiosity becomes :

$$B(s,t) \approx \sum_{k=1}^{N'} b_k^i \Gamma_k(s,t) \tag{3}$$

where $b_k^i$ is the coefficient associated to the $k^{th}$ function of the base.

For each surface $i$, $b_k^i$ coefficients are obtained by substituting (3), and respectively expressions of exitance and reflectivity, in equation (1) and by using the classical inner product between the radiosity $B_i(s,t)$ and the $k^{th}$ function of the base. We obtain :

$$b_k^i = e_k^i + \sum_{j=1}^{N} \sum_{l=1}^{N'} b_j^l \left( \int K_{ij}(s,t,u,v) \Gamma_l(u,v) \, du \, dv \right) \tag{4}$$

and finally :

$$b_k^i = e_k^i + \sum_{j=1}^{N} \sum_{l=1}^{N'} b_j^l K_{ij}^{kl} \tag{5}$$

with

$$K_{ij}^{kl} = \left( \int K_{ij}(s,t,u,v) \Gamma_l(u,v) \, du \, dv \right) \tag{6}$$

This coefficient represents a kind of generalised form factor expressing the energy exchanged between the $k^{th}$ function associated with surface $i$ and the $l^{th}$ function associated with surface $j$. Methods for computing the $K_{ij}^{kl}$ terms use either traditional rules of quadratic integration [30], or Monte Carlo’s techniques [18]. Equation (5), where the unknown are the radiosity coefficients $b_k^i$ for each surface, can be solved by using indifferently a traditional direct numerical method, or via any progressive refinement techniques [7]. Indeed, since the surface indices $i,j$ and function indices $k,l$ are independent of each other, (5) is still a linear equation. Moreover, in the Galerkin method, shadow is moving out of the light calculation (visibility $VIS$ is equal to 1). Zatz uses shadow masks but other shadow algorithms [6, 9, 16, 24, 25, 26] could be used.

3. Our approach

In this section, we present a new algorithm for producing a large animation sequence with global illumination determination. Our main goal is to compute at time $T + \Delta T$, the new illumination of a scene knowing the whole radiosity of the previous frame, at time $T$. It leads to a new progressive algorithm that computes only radiosity variations with new approximated form factors. It can manage any kind of surface -even light source- in any complex translation.

In the first time, we analyze the differences between two successive frames. Then we focus on the determination of the new form factor and the mean to obtain them quickly. Then we discuss some discontinuity situations which deserve special treatments. Finally, we set our algorithm used to compute animation sequence.

3.1. Radiosity modifications between two frames

To simplify, let us consider the contribution of one moving surface $j$ to surface $i$ as shown in figure 2. Time dependent values are representing by adding a quote at time $T+\Delta T$. 

![Figure 2. Radiosity modifications between two frames](image)
The fundamental idea is to find a mathematical relation between $b_i^h$ and $b_j^k$. Expressing equation (1) for the two surfaces $i$ and $j$, the radiosity difference is:

$$B_j(s,t)-B_i(s,t) = E_j(s,t)-E_i(s,t) + \int_{u,v} [K_{ij} B_j - K_{ij} B_i](s,t,u,v) \, du \, dv \quad (7)$$

Using notation $K'_{ij} = K_{ij} + \Delta K_{ij}$, $B'_i = B_i + \Delta B_i$, $E'_i = E_i + \Delta E_i$ equation (7) could be reformulated:

$$\Delta B_i = \Delta E_i + \int_{u,v} [\Delta K_{ij} \cdot B_j + \Delta B_j \cdot (K_{ij} + \Delta K_{ij})] \quad (8)$$

If we develop the radiosity $B_j$ by substituting (3) in equation (8) and projecting the result onto the $k^h$ base function, a relation is obtained:

$$\Delta b_i^k = \Delta e_i^k + \sum_{i=1}^{N} b_j^k \Delta K_{ij} + \sum_{i=1}^{N} \Delta b_j^k \cdot (K_{ij} + \Delta K_{ij}) \quad (9)$$

with

$$\Delta K_{ij} = \left( \int_{u,v} (K'_{ij} - K_{ij})(s,t,u,v) \Gamma_i(u,v) \mid \Gamma_k(s,t) \right) \quad (10)$$

We should notice that equation (9) can be reversed to determine also the modification in illumination of the moving surface $j$ - created by its own move - from static surface $i$. We just have to consider that surface $i$ has a virtual relative move opposite to the original move of surface $j$ and compute the coefficient $\Delta K_{ji}$.

### 3.2. Calculation of coefficients $\Delta K$

Henceforth, the main problem is to determine for each frame the new coefficients $\Delta K$. If we consider that the surface $j$ follows a translation movement in a direction $\vec{p}_0$ as shown in figure 3, the idea is to obtain, with a limited development in $p$ of the expression $\Delta K$, a polynomial expression of $\Delta K$ like:

$$\Delta K = \sum_{n=1}^{\infty} p^n \phi_n$$

where $\phi_n$ depend only on $\vec{p}_0$.

If, from an initial reference position, each coefficient $\phi$ is computed, then for each displacement $p$ of the surface $j$ in the direction $\vec{p}_0$, $\Delta K$ will be efficiently computed.

We start with the expression of $K$ in equation (2) with terms:

$$F_{i\rightarrow j}(s,t,u,v) = - \frac{n_d \cdot \vec{n}_i \cdot \vec{p}_0}{\pi r^2} \hbox{ and } A_j(u,v) = \left\| \frac{\partial \vec{N}}{\partial u} \wedge \frac{\partial \vec{N}}{\partial v} \right\|$$

So we have:

$$K_{ij} = -\frac{\rho_i \cdot n_d \cdot \vec{n}_i \cdot \vec{p}_0}{\pi r^2} \left\| \frac{\partial \vec{N}}{\partial u} \wedge \frac{\partial \vec{N}}{\partial v} \right\| \quad (11)$$

Let consider the expression of $r^n$:

$$r^n = (r + \vec{p})^n = r^n \left( 1 + 2 \frac{\vec{p}}{r^2} + \frac{\vec{p}^2}{r^2} \right)^n$$

We denote $a = \vec{p}_0 \cdot \vec{r}$ and $\beta = \vec{p}_0^2$. The limited development of the inverse of equation (12) gives:

$$\frac{1}{r^n} \sum_{n=0}^{\infty} (-1)^n (n+1) \left( 2 \frac{\vec{p}}{r} + \frac{\vec{p}^2}{r^2} \right)^n$$

and is defined only if:

$$r > (1 + \sqrt{2}) \cdot p$$

Using the following notation:

$$\begin{cases} a = (\vec{n}_d \cdot \vec{r}) \cdot (\vec{n}_s \cdot \vec{x}) \\ b = (\vec{n}_d \cdot \vec{p}_0) \cdot (\vec{n}_s \cdot \vec{x}) + (\vec{n}_d \cdot \vec{r}) \cdot (\vec{n}_s \cdot \vec{p}_0) \\ c = (\vec{n}_d \cdot \vec{p}_0) \cdot (\vec{n}_s \cdot \vec{p}_0) \end{cases}$$

Figure 3. Details of a moving surface and a static surface
we have

\[ K'_{ij} = -\frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a + b p + c p^2 \right), \]

\[ \sum_{n=0}^{\infty} (-1)^n (n+1) \left( 2 \frac{\alpha}{r^2 p} + \beta \frac{p^2}{r^2} \right)^n \]  
(14)

or

\[ K'_{ij} = K_{ij} - \frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a + b p + c p^2 \right), \]

\[ -\frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a + b p + c p^2 \right), \]

\[ \sum_{n=1}^{\infty} (-1)^n (n+1) \left( 2 \frac{\alpha}{r^2 p} + \beta \frac{p^2}{r^2} \right)^n \]

Finally \( p \) is extracted from the last term of the preceding equation:

\[ \sum_{n=1}^{\infty} \frac{(-1)^n (n+1)}{p^{2n}} \sum_{k=0}^{n} \binom{n}{k} (2\alpha p)^{n-k} p^k \]

\[ = \sum_{n=1}^{\infty} \frac{(-1)^n (n+1)}{p^{2n}} \sum_{k=0}^{n} \binom{n}{k} (2\alpha)^{n-k} p^{n+k} \]

\[ = \sum_{n=1}^{\infty} \frac{(-1)^n (n+1)}{p^{2n}} \sum_{q=0}^{n} \binom{n}{q} (2\alpha)^{2n-q} \beta^q p^q \]

\[ \text{avec } q = n + k \]

\[ = \sum_{n=1}^{\infty} \frac{(-1)^n (n+1)}{p^{2n}} \sum_{q=0}^{n} \binom{n}{q} (2\alpha)^{2n-q} \beta^q p^q \]

\[ = \sum_{n=1}^{\infty} \frac{(-1)^n (n+1)}{p^{2n}} \sum_{q=0}^{n} \binom{n}{q} (2\alpha)^{2n-q} \beta^q p^q \]

\[ = \sum_{p \geq 1} \xi_q \]

and

\[ (a + b p + c p^2), \sum_{q \geq 1} p^q \xi_q \]

\[ = a \sum_{q \geq 1} p^q \xi_q + b \sum_{q \geq 1} p^{q+1} \xi_q + c \sum_{q \geq 1} p^{q+2} \xi_q \]

\[ = a \xi_1 + b \xi_2 p + c \xi_3 p^2 + \sum_{q \geq 3} p^q (a \xi_q + b \xi_{q-1} + c \xi_{q-2}) \]

We have then

\[ K'_{ij} - K_{ij} = \sum_{n \geq 1} p^n \phi_n^{ij}(s, t, u, v, \rho_0) \]  
(15)

with

\[ \phi_n^{ij} = -\frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a \xi_1 + b \right) \]

\[ \phi_n^{ij} = -\frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a \xi_2 + b \xi_1 + c \right) \]

\[ \forall n \geq 3, \phi_n^{ij} = -\frac{\rho_i}{\pi r^4} \left[ \frac{\delta N}{\delta u} \right] \left[ \frac{\delta N}{\delta v} \right] . \left( a \xi_n + b \xi_{n-1} + c \xi_{n-2} \right) \]

\[ \forall n, \xi_n = \sum_{q=n}^{n} p^q \left( \frac{\alpha}{\beta} \right)^q \]

And finally,

\[ \Delta K_{ij} = \sum_{n \geq 1} p^n \left( \int_{u,v} \phi_n^{ij}(s, t, u, v) \right) \Gamma_k(u,v) | \Gamma_k(s,t) \]  
(16)

Equation (16) allows to compute efficiently coefficients \( \Delta K \). Indeed, coefficients \( \phi \) are recursively defined and depend only on the direction \( \rho_0 \), on the degree of approximation \( n \), and on both surfaces \( i,j \) and their function \( k,l \). So coefficients \( \phi \) could be computed in advance for any elementary directions, like axes \( \bar{x}, \bar{y} \) and \( \bar{z} \), and used for all move \( p \) in these directions.

### 3.3. Condition of validation and discontinuity

The equation (16) is usable only if condition (13) is satisfied and if discontinuities have not occurred because continuity is an implicit condition of all previous calculations.

Discontinuities occur when a receiving surface, or a part of it, enters or goes out of the visibility of the moving surface. For discontinuities, we have to check if receiving surface enters or goes out of the positive half space defined by the emitter, as shown in figure 4. Simple geometrical computations, like BSP Tree [6], can determine such particular situations.

![Figure 4. Example of positive and negative half space of an emitting surface](image)

When discontinuity occurs, we have to compute the new coefficients \( K' \) explicitly using equation (6). At this time, we get the correct value of \( K \) and errors, involved by multiple addition of \( \Delta K \), disappear.
The condition (13) expresses that, for a particular elementary direction $\mathbf{p}$, our approximation is defined near the reference point where we compute the limited development and the coefficients $\varphi$. $r$ represents the distance between moving surface and static surface at this reference position and $p$ the norm of the path of the moving surface from it. So these two values are stored for each direction and each receiving surface, and updated to control condition (13).

When this condition is no more satisfied in a particular direction, we have to compute explicitly $K^i$ and also the coefficients $\varphi$. $r$ and $p$ are also updated (always for that direction). Once computed, the new coefficients $\varphi$ remain valid until the new condition (13) is not satisfied i.e. moving surface moves too far in that direction.

3.4. Illumination of the scene for a new frame

The following system of equation allows to compute incrementally the global illumination solution of a scene:

$$\Delta b^k_i = \Delta e^k_i + \sum_{j=1}^{N'} b^j_i \Delta K^k_{ij} + \sum_{i=1}^{N'} \Delta b^j_i (K^k_{ij} + \Delta K^k_{ij})$$

with

$$\Delta K^k_{ij} = \left( \left[ \int_{a,v} (K^i_{ij} - K_{ij}) (u,v) \Gamma_i(u) \right| \Gamma_k(v) \right)$$

This system is equivalent to the galerkin radiosity equation (5) and can be seen as a progressive equation with the unknown variables $\Delta b$. There are only two differences:

- The emission term has been replaced by an eventual change in exitance $\Delta e^k_i$, plus a sum $b^j_i \Delta K^k_{ij}$ that represents the illumination difference involved by the move.

- The unknown variables have changed from $b_i$ to $\Delta b_i$.

Every complex path, like splines, could be decompose in a sequence of little moves along elementary directions (the three axes for example). Then for each new image, we determine directions and step of the moving surfaces (we don’t have to know the movement in advance). Depending on condition (13) and discontinuity, we use equations (16) or (6) to compute the two first terms of equation (9). Computing these terms initialised the progressive refinement algorithm, that we use to propagate the light in the scene with the new form factors. When all the light have been propagated, we have the illumination of the entire scene without shadow effects. To represent them, we use the algorithm of the Shadow Volume Binary SSpace Tree (SVBSP Tree) [6]. We choose this algorithm for its simplicity and its efficiency. Indeed, hard shadows are obtained very quickly and the realism generated by the Galerkin method is enough to obtain visually correct results. Other shadow algorithms [9, 16, 25] could be used.

Aliasing appears in the animation sequence but few modifications allow to erase it. Aliasing is due to discretisation of the scene and of the radiosity solution (and especially shadows). Jittering avoids the first kind of aliasing and takes only one second (or less with graphic cards allowing accumulation buffer). The radiosity solution over a surface is represented with a regular grid which involve also aliasing. We can use a finer grid or jitter it. Memory cost will limit the first approach but will be quicker. Jittering the grid forces redrawing the scene four or eight times, so will take about 1 or 2 seconds per image.

4. Results

The main computation time in the galerkin algorithm, except for shadow mask, is due to the determination of pseudo form factor. Our algorithm avoids computing them for each frame. In fact we made a costly computation in the beginning to obtain coefficients $\varphi$ but then for each image of the sequence we don’t have to compute explicitly the new form factors. Moreover, between two positions, we only compute the difference of energy involved by the move. Indeed, we calculate what is really necessary since we use a progressive method for each position (normal progressive method) and also between each position (our temporal progressive method). Thus, we can manage complex dynamic scene easily and simply since the main part of these scenes will probably not be concerned by the energy exchanges involved by the move (like the bressenham algorithm). Notice also that any complex path could be represented if it can be decomposed in little elementary moves. And we don’t have to know the move in advance.

We present here some results and images from animation sequences computed with a 500 MHz processor and a common PC graphic card. The first scene, illustrated in figure 5.a, is a simple scene with a blue surface passing in front of a light and a second "ceiling light" moving left and right in translation. The right wall is correctly lighted in blue when the blue occluder enters just in front of the light. The red wall in the left receives also more light when the ceiling light comes nearest. Table 1 shows computation times and benefits of our method. We shall point out that this scene constitutes a worst case because the blue panel is very close to the light (and so $r$ is small) and it enters progressively in the light positive half space.

<table>
<thead>
<tr>
<th>100 frames of figure 5.a</th>
<th>Computation of $\varphi$</th>
<th>Time / frame</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galerkin progressive method</td>
<td>/</td>
<td>4.02 s</td>
<td>402 s</td>
</tr>
<tr>
<td>Our method</td>
<td>18.59 s</td>
<td>1.41 s</td>
<td>160 s</td>
</tr>
</tbody>
</table>

Table 1. Result for scene of figure 5.a
Figure 5. a. Simple Scene b. Complex scene: The billiard

The second scene, shown in figure 5.b, is a complex scene with a moving light. We show in table 2 computation times for that particular scene. Despite multiple discontinuities that occur, timesaving remains the same. We also present in the animation movies a more complicated scene (figure 6), a cathedral interior, with 140 surfaces representing about 110,000 polygons. Time saving is still equivalent to the previous scene. We should notice that we are able to represent any complex move in translation unlike what is showed in the animations.

<table>
<thead>
<tr>
<th>160 frames of figure 5.b</th>
<th>Computation of $g$</th>
<th>Time / frame without Stree</th>
<th>Time / frame</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galerkin method</td>
<td>2.93 s</td>
<td>2.03 s</td>
<td>469 s</td>
<td></td>
</tr>
<tr>
<td>Our method</td>
<td>18.59 s</td>
<td>1.13 s</td>
<td>0.24 s</td>
<td>185 s</td>
</tr>
<tr>
<td>Benefits (%)</td>
<td>61 %</td>
<td>88 %</td>
<td>60 %</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Result for scene of figure 5.b

5. Conclusion and perspective

In this paper, we have presented a new algorithm that allows to solve the case of any surface in translation—even light source—in the radiosity method. Timesaving is obtained by avoiding the computation in each frame of the sequence, for each moving object, the form factors associated with the mobile elements. Moreover, this algorithm makes only the necessary calculations what gives us a very efficient time for each new frame computation. Perspectives relate to shadows. We concentrated only on hard shadow but other algorithms could be used to represent soft shadows. As the Galerkin method does not require any discretisation of the scene, the storage cost of information of visibility can be managed. Another subject of development relates to the rotation effects of an emitting surface. The problem is more complicated since a slight rotational movement can involve major change in the illumination. Nevertheless, with appropriate limitations, for small oscillations, we can obtain a correct and useful approximation.

Figure 6. the cathedrale scene

References
