AN EMPIRICAL EVALUATION OF A GPU RADIOSITY SOLVER

Guenter Wallner
Institute for Art and Technology, University of Applied Arts, Oskar Kokoschka Platz 2, Vienna, Austria

Keywords: Global illumination, Radiosity, GPU programming, Projections.

Abstract: This paper presents an empirical evaluation of a GPU radiosity solver which was described in the authors previous work. The implementation is evaluated in regard to rendering times in comparison with a classical CPU implementation. Results show that the GPU implementation outperforms the CPU algorithm in most cases, most importantly, in cases where the number of radiosity elements is high. Furthermore, the impact of the projection – which is used for determining the visibility – on the quality of the rendering is assessed. Results gained with a hemispherical projection performed in a vertex shader and with a real non-linear hemispherical projection are compared against the results of the hemicube method. Based on the results of the evaluation, possible improvements for further research are pointed out.

1 INTRODUCTION

Radiosity has been and is still a widely used technique to simulate diffuse interreflections in three-dimensional scenes. Radiosity methods were first developed in the field of heat transfer and in 1984 generalized for computer graphics by (Goral et al., 1984). Since then, many variations of the original formulation have been proposed, leading to a rich body of literature available on this topic. See e.g. the book by (Cohen and Wallace, 1995) for a good overview of the field. An extensive comparison of different radiosity algorithms can be found in (Willmott and Heckbert, 1997).

In recent years research on GPU accelerated radiosity algorithms intensified. (Nielsen and Christensen, 2002) used hardware texture mapping to accelerate the hemicube method and (Carr et al., 2003) were the first to utilize floating point textures to store the radiosity values. (Coombe et al., 2003; Coombe and Harris, 2005) described the first radiosity solver for planar quadrilaterals that performed all steps on programmable graphics hardware. In (Wallner, 2008) we described a GPU radiosity system for arbitrary triangular meshes which built upon the work mentioned above.

The purpose of this paper is twofold. First, presenting the results of an empirical evaluation of our GPU radiosity solver in regard to rendering times and on how the projection used for creating the visibility map effects the final image quality and secondly to point out possible improvements for further development.

The reminder of this paper is structured as follows. Section 2 gives a short overview of the various steps of the radiosity implementation. In Section 3 timings for the individual steps are given for a set of sample scenes. Section 4 compares rendering times with a CPU implementation and the impact of the applied projection on the quality of the result is discussed in Section 5. The paper is concluded in Section 6.

2 OVERVIEW

This section gives a short overview of the main steps performed by the progressive radiosity solver, which are as follows:

Preprocess. In the preprocess two textures – which store radiosity and residual energy values – are generated for each triangle which themselves are placed in larger lightmaps to reduce texture switching later in the process. Furthermore the required framebuffers, auxiliary textures and other OpenGL related resources are allocated.

Next Shooter Selection. In this step a mipmap pyramid is constructed for each residual lightmap until each triangle is represented by a single texel. These values are read back to the CPU and the triangle with the highest value is selected as next shooter. This step is performed using a ping-pong technique. Standard
OpenGL mipmapping cannot be used because texels not occupied by the triangle have to be omitted. For the shooter, four textures — world position map, normal map, color map and intensity map — are generated which are necessary for calculation of the energy transport. Shooting is performed from a user-defined mipmap level of these textures. This way the accuracy and speed of the calculation can be influenced.

Render Visibility Texture. To determine the visible surfaces a hemicube projection, a stereographic projection (performed in the vertex shader) or the nonlinear hemisphere projection by Gascuel et al. (Gascuel et al., 2008) is used to generate a depth texture from the shooter’s viewpoint.

Adaptive Subdivision. If adaptive subdivision is enabled, occlusion queries are used during rendering of the visibility texture, to check on which triangles shadow boundaries are located. If the ratio between visible texels and texels in shadow is below certain thresholds and the maximum subdivision level is not reached, the triangle is subdivided.

Update Receivers. Visible or partially visible triangles (determined by checking the outcome of an occlusion query) are back-projected into the visibility texture and their depth values are compared against the stored values. If they match, then the radiosity and residual maps are updated accordingly. Calculation of the form factors is therefore performed in this step.

Postprocess. Because the rasterized triangles cannot be directly used for texturing - polygon rasterization rules differ from texture sampling rules (Segal and Akeley, 2003) — missing values are linearly interpolated in a shader. This step is performed once after the progressive radiosity solver has finished.

The system was implemented in C++ and OpenGL with the shaders written in Cg. For a more in-depth description of the solver refer to (Wallner, 2008) and (Wallner, 2009). In the remainder of this paper we will refer to one cycle of the progressive radiosity solver as an iteration in contrast to a shooting step from a single texel of the mipmap texture of the selected shooting triangle.

3 PERFORMANCE

To assess the time spent on the above mentioned steps we used four different test scenes, which vary in the number of triangles and light sources. All the measurements in this work were taken on an Intel Core2 CPU with 2.13 GHz, 3.5 GB RAM with a Geforce 8800GTS with 640MB DDR3 Ram. The scenes (except the Cornell Box) are shown in Figure 2 whereas Figure 1 depicts the measurements in case of the nonlinear hemisphere. The hemicube is rendered with a single-pass cubemap geometry shader with multiple attached layers. This has the advantage that an occlusion query (required for subdivision and visibility purposes) counts all rasterized fragments over all attached layers, which makes management of them easier than with a multi-pass approach. Performing the hemispherical projection in the vertex shader reduces the time needed for the creation of the visibility texture by about 30 to 55 percent. However, if the scene is not carefully triangulated the hemispherical projection can lead to severe artifacts (see Section 5).

Time for rendering the visibility texture depends solely on the number of triangles, whereas the time needed for the next shooter selection depends on the number of lightmaps and the ratio of the lightmap size to the stored texture size. Updating the receivers is by far the most time consuming function in the process. In this regard the biggest restriction is currently that GPUs do not allow to simultaneously read from and write to the same position in a texture. Therefore it is necessary to render the new radiosity and residual energy to intermediate textures and copy them afterward to the correct position in the lightmap. Another possibility would be to maintain two textures for storing the radiosity and further two for the residual energy and use them in a ping pong technique. We rejected this method because it would double the memory requirements.

The cost for the post-process primarily depends on the number of triangles. To reduce the required time
for the post-process we initially hoped that with con-
servative rasterization (Hasselgren et al., 2005) the
textures could be directly used for nearest-neighbor
texture sampling. This, however, proved to be diffi-
cult since the geometry itself is inflated and therefore
the interpolation of normals and other attributes is al-
tered, which inevitable leads to discontinuities at the
boundaries of the triangles. It is clear that the cost of
the pre- and post-process amortizes with increasing
number of performed iterations.

4 CPU COMPARISON

To compare the performance of our solution with
a software implementation we chose the radiosity
solver of the open source software Blender 2.48a
(Blender Foundation, b). We opted for Blender be-
cause it is freely available and the settings of its ra-
diosity system can be closely matched with the ones
from our implementation. The hemicube resolution
was set to 512 × 512 and in case of the stereographic
projection a 2048 × 2048 resolution was used. The
calculation was terminated after a fix number of iter-
ations. We have preferred iterations over convergence
because the time needed to reach a certain conver-
gence depends on the mipmap level chosen for shoot-
ing in our method. Ensuring the same number of it-
ervations between the CPU and GPU implementation
would therefore be hard to control. Adaptive sub-
division was disabled in our implementation, since
Blender performs no adaptive refinement (Blender
Foundation, a). Table 1 lists the timings for differ-
ent scenes. For the test scenes our GPU implementa-
tion achieved speed-ups of up to 46 times. In con-
trast to the CPU implementation, the calculation time
is nearly unaffected by the number of elements (that
is, the resolution of the radiosity maps). Calculation
time on the GPU is mostly determined by the num-
ber of triangles in the scene, since for each triangle a
texture copy is performed during update of the radios-
ity and residual textures. The non-linear hemisphere
projection achieved about 20 to 40 percent faster ren-
dering times than the hemicube method.

In cases where the number of elements is very low
the CPU outperforms the GPU, as it is e.g. the case
in the low resolution temple scene. In such scenarios
the cost of the initial pre-process of transferring the
data to the GPU, the allocation of resources and the
required post-process are high compared to the actual
processing time.

5 PROJECTIONS

In this section the influence of the projection utilized
for creating the visibility map on the visual quality of
the result is discussed. The classical method for deter-
mining visibility from a given point is the hemicube,
as proposed by (Cohen and Greenberg, 1985). The
drawback is that the scene has to be rendering five
times to project all possible receiving surfaces onto
the individual sides of the hemicube. Since visibility
determination is a frequently performed task in the
radiosity process, it is beneficial to reduce rendering
time. (Beran-Koehn and Pavicic, 1991; Beran-Koehn
and Pavicic, 1992) used a cubic tetrahedron instead a
hemicube, thereby reducing the number of projecting
planes to three. The number of projections were fur-
ther reduced by using a single-plane projection (e.g.
(Sillion and Puech, 1989)) or hemisphere (base) pro-
jections which, to our knowledge, were first proposed
by (Spencer, 1992). In this case the projection is a
simple normalization process and clipping can sim-
ply be performed against the base plane of the hemi-
sphere. To ensure an accurate projection, (Spencer,
1992) calculates the degree of the arc between pairs of
points projected onto the hemisphere basis and uses
this value to calculate intermediate points along the
edge. A scan-conversion algorithm is then used to
Table 1: Performance comparison between a CPU radiosity solver \(t_{\text{cpu}}\) and our implementation \(t_{\text{gpu,cube}}\) for the hemicube and \(t_{\text{gpu,sphere}}\) for the non-linear hemispherical projection with the speed-up in brackets. The number of elements is given in the columns \(n_{\text{cpu}}\) and \(n_{\text{gpu}}\). Because our method uses power-of-two textures the number of elements differ, nonetheless we tried to match them as closely as possible. \(n_t\) lists the number of triangles and \(n_s\) the number of shootings.

<table>
<thead>
<tr>
<th>Scene</th>
<th>(s)</th>
<th>(n_t)</th>
<th>(n_{\text{cpu}})</th>
<th>(n_{\text{gpu}})</th>
<th>(t_{\text{cpu}}) [\text{s}]</th>
<th>(t_{\text{gpu,sphere}}) [\text{s}]</th>
<th>(t_{\text{gpu,cube}}) [\text{s}]</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbox_low</td>
<td>32</td>
<td>258</td>
<td>9248</td>
<td>11802</td>
<td>1.97</td>
<td>0.79 ((\times 2.5))</td>
<td>1.15 ((\times 1.7))</td>
</tr>
<tr>
<td>cbox_medium</td>
<td>32</td>
<td>258</td>
<td>82976</td>
<td>94750</td>
<td>5.69</td>
<td>0.75 ((\times 7.6))</td>
<td>1.13 ((\times 5.0))</td>
</tr>
<tr>
<td>cbox_high</td>
<td>32</td>
<td>258</td>
<td>1026464</td>
<td>1115126</td>
<td>54.19</td>
<td>1.17 ((\times 46.3))</td>
<td>1.48 ((\times 36.6))</td>
</tr>
<tr>
<td>temple_low</td>
<td>24</td>
<td>15872</td>
<td>117248</td>
<td>126078</td>
<td>5.60</td>
<td>9.74 ((\times 0.6))</td>
<td>12.58 ((\times 0.4))</td>
</tr>
<tr>
<td>temple_medium</td>
<td>24</td>
<td>15872</td>
<td>507320</td>
<td>516312</td>
<td>18.27</td>
<td>9.92 ((\times 1.8))</td>
<td>12.72 ((\times 1.4))</td>
</tr>
<tr>
<td>temple_high</td>
<td>24</td>
<td>15872</td>
<td>1741208</td>
<td>2074242</td>
<td>53.13</td>
<td>10.03 ((\times 5.3))</td>
<td>13.18 ((\times 4.0))</td>
</tr>
<tr>
<td>office_low</td>
<td>96</td>
<td>21704</td>
<td>264081</td>
<td>254572</td>
<td>36.17</td>
<td>18.46 ((\times 1.3))</td>
<td>42.46 ((\times 0.8))</td>
</tr>
<tr>
<td>office_medium</td>
<td>96</td>
<td>21704</td>
<td>747375</td>
<td>854341</td>
<td>84.62</td>
<td>18.03 ((\times 4.7))</td>
<td>42.44 ((\times 2.0))</td>
</tr>
<tr>
<td>office_high</td>
<td>96</td>
<td>21704</td>
<td>1846137</td>
<td>1843771</td>
<td>202.54</td>
<td>18.73 ((\times 10.8))</td>
<td>43.03 ((\times 4.7))</td>
</tr>
<tr>
<td>living_low</td>
<td>32</td>
<td>38147</td>
<td>474744</td>
<td>472317</td>
<td>37.23</td>
<td>20.18 ((\times 1.8))</td>
<td>33.10 ((\times 1.1))</td>
</tr>
<tr>
<td>living_medium</td>
<td>32</td>
<td>38147</td>
<td>861540</td>
<td>866330</td>
<td>44.41</td>
<td>20.05 ((\times 2.2))</td>
<td>33.24 ((\times 1.3))</td>
</tr>
<tr>
<td>living_high</td>
<td>32</td>
<td>38147</td>
<td>1570884</td>
<td>1612067</td>
<td>73.88</td>
<td>20.61 ((\times 3.6))</td>
<td>33.45 ((\times 2.2))</td>
</tr>
</tbody>
</table>

Figure 3: Left: Incorrect shadows as the result of erroneous visibility information due to the hemispherical projection. Middle: Adaptively tessellating the surface around the shadow boundary reduces the errors enormously. Right: Solution of the hemicube method on the original mesh.

determine the area covered by the projected surface element. The method, however, fails to remove hidden surfaces in certain cases, which was resolved by (Doi and Takayuki, 1998) by subdividing the surface – based on a solid angle criterion – into smaller triangles instead of using edge-subdivision. Hemisphere projections are currently frequently performed in vertex shaders on programmable GPUs, e.g. (Coombe et al., 2003; Barsi and Jakab, 2004; Coombe and Harris, 2005). However, since only vertices are affected by the hemispherical projection – straight edges are mapped to straight edges instead of elliptical segments – errors are introduced. To approximate the correct curvature, one solution would be to tessellate the surfaces finely enough, as e.g. done in (Coombe et al., 2003). This increases the triangle count of the scene considerably and consequently rendering times. Assuming that no two surfaces overlap, it is sufficient to perform the back-projection in the vertex shader as done in (Wallner, 2008; Barsi and Jakab, 2004). However, if surfaces overlap, the relative position information between the surfaces is distorted in the projection which results in more or less incorrect shadow boundaries. Figure 3 illustrates this problem. However, adaptive subdivision methods (see e.g. (Cohen and Wallace, 1995)) to improve the accuracy of the radiosity solution in areas of high-frequency lighting also reduce the errors due to the hemisphere projection significantly. Therefore it seems to be unnecessary to uniformly tessellate the mesh just for the purpose of the projection.

A further problem in conjunction with too coarse meshes is that the projection of some triangles may be extremely deformed and therefore are not rendered at all, although they should be. These triangles are misleadingly considered not visible and thus omitted during the radiosity exchange. Figure 4 shows an example where this is the case. This problem is intrinsic to hemispherical projections because distortions increase for directions further from the viewing direction. Using a paraboloid projection would reduce these errors since the sampling rate is more uniform than in the spherical case (see (Heidrich and Seidel, 1998)). However, the actual problem remains.

(Kautz et al., 2004) proposed a method for spherical rasterization, by defining a plane through each edge of the triangle and the projection center. Then a pixel is considered inside the projected image of the
the RMS is given by
\[ \text{RMS} = \sqrt{\frac{1}{m \cdot N \cdot X \cdot Y} \sum_{n=1}^{N} \sum_{x=1}^{X} \sum_{y=1}^{Y} |v^n_{xy}| \cdot a^n_{xy}} \] (1)

where \(v^n_{xy}\) is the color-difference vector between the \(R, G,\) and \(B\) values at pixel position \((x, y)\) between the \(n\)-th lightmap produced by hemispherical projection and the corresponding lightmap computed with the hemicube method.

Pixels which do not belong to a triangle in the lightmap are neglected by multiplication with the the alpha-value \(a^n_{xy}\) at the current position, which is either 1 or 0 if the pixel is occupied by a triangle or not. This is necessary because such pixels are not part of the solution and would only dilute the RMS value. \(m\) is therefore the number of occupied pixels \((a = 1)\) in all lightmaps. Since the result of the adaptive subdivision can vary between the hemicube and hemisphere solution, the result of the adaptive subdivision gained by the hemisphere solution was used for the hemicube method. Table 2 lists these root mean square values for two different scenes. In this case we did not keep the iterations fixed but rather aborted the calculation once a given convergence threshold was reached. The measurements were performed for different subdivision levels with different texture sizes. For the original mesh \(256 \times 256\) textures were used, which yields a resolution of \(512 \times 512\) for a quad. This means the regular subdivision with level 3 and a texture size of 64 matches the original resolution most closely.

First of all, the RMS-error in case of the Cornell Box is higher throughout than for the column scene since the light source is much closer to the scene geometry. As is apparent from Table 2 (and as expected) a regular subdivision reduces the RMS error quite significantly, up to 13 times. This is the case even if the overall resolution over a non-subdivided quad is smaller than the \(512 \times 512\) texels used on the original geometry (e.g. using a regular subdivision with level 3 and texture size 32 yields a better result than achieved with the original mesh). However, a regularly subdivided mesh also increases rendering times considerably. This is, on the one hand, due to the increased number of triangles and, on the other hand – which may not be obvious at first – due to the fact that more triangles covering the same area means less energy per triangle which in turn requires more iterations to reach a given convergence threshold.

We are looking into two options to further decrease rendering times:

1. Currently a single texture size is used for the whole mesh (although sizes can vary between meshes), which may be suboptimal since triangles sizes can vary over the mesh. In this sense, it may be promising to base the texture resolution on triangle size. However, first results showed interpolation artifacts along triangle edges due to the different resolutions.

2. The mipmap level used for shooting may be decided adaptively rather than being fixed for the whole calculation. The level may also depend on the average energy of the current shooting tri-
Table 2: RMS error for the hemispherical projection (compared to the hemicube solution). The table lists the type of subdivision \(\text{type}_{sub}\) which is either no (no subdivision), adapt (adaptive subdivision) or reg (regular refinement), and the level of subdivision. The texture size is given in column size and the rendering time is listed in column \(t\). If a solution is considered perceptually indistinguishable from the hemicube solution it is highlighted in green, otherwise in red.

<table>
<thead>
<tr>
<th>(\text{type}_{sub})</th>
<th>level</th>
<th>size</th>
<th>Columns</th>
<th>(t[s])</th>
<th>(RMS)</th>
<th>Columns</th>
<th>(t[s])</th>
<th>(RMS)</th>
<th>Cornell Box</th>
<th>(t[s])</th>
<th>(RMS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>0</td>
<td>256</td>
<td>hemisphere</td>
<td>1.836580</td>
<td>0.000279</td>
<td>non-linear hemisphere</td>
<td>2.191166</td>
<td>0.005399</td>
<td>hemisphere</td>
<td>1.543515</td>
<td>0.001286</td>
</tr>
<tr>
<td>adapt</td>
<td>3</td>
<td>32</td>
<td>0.0007547 0.997896</td>
<td>0.000953</td>
<td>2.134127</td>
<td>0.001772</td>
<td>14.641043</td>
<td>0.000458</td>
<td>20.789242</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reg</td>
<td>3</td>
<td>64</td>
<td>0.0000331 9.142668</td>
<td>0.000104</td>
<td>10.184814</td>
<td>0.002147</td>
<td>5.382289</td>
<td>0.000278</td>
<td>4.373851</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adapt</td>
<td>3</td>
<td>64</td>
<td>0.000789 1.179061</td>
<td>0.000397</td>
<td>2.295658</td>
<td>0.000766</td>
<td>79.602470</td>
<td>0.000231</td>
<td>37.284330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reg</td>
<td>3</td>
<td>64</td>
<td>0.000133 28.540912</td>
<td>0.000057</td>
<td>29.953424</td>
<td>0.000147</td>
<td>5.382289</td>
<td>0.000278</td>
<td>4.373851</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adapt</td>
<td>2</td>
<td>32</td>
<td>0.003907 0.903491</td>
<td>0.001434</td>
<td>1.968630</td>
<td>0.040701</td>
<td>1.46952</td>
<td>0.011807</td>
<td>3.244810</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reg</td>
<td>2</td>
<td>64</td>
<td>0.0001287 2.791759</td>
<td>0.000457</td>
<td>3.137670</td>
<td>0.004473</td>
<td>2.77055</td>
<td>0.000194</td>
<td>4.120059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adapt</td>
<td>2</td>
<td>64</td>
<td>0.0001726 1.004140</td>
<td>0.000576</td>
<td>2.064948</td>
<td>0.020571</td>
<td>1.644687</td>
<td>0.004213</td>
<td>5.477550</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reg</td>
<td>2</td>
<td>64</td>
<td>0.000591 3.272110</td>
<td>0.000187</td>
<td>3.731351</td>
<td>0.001784</td>
<td>4.809930</td>
<td>0.000709</td>
<td>7.562907</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This would make sense, since triangles with less residual energy make a smaller contribution to the overall image. In the same way the most noticeable contributions are being made at the beginning of the progressive radiosity algorithm.

However, as evident from Table 2, the RMS error can be already reduced by using adaptive subdivision which simultaneously keeps rendering times low. In fact, it reduces the most noticeable errors near shadow boundaries as depicted in Figure 3. In some cases rendering times are even lower than with the original geometry. This is the case because lower resolutions were used, meaning that there are less shootings from a single triangle. In contrast to the regular subdivision, areas with no shadow boundaries (which therefore have higher residual energies on average) are not refined in the same extent which in turn means that the energy is not distributed over that many triangles and therefore the number of iterations is not increased by the same extent as is the case for regular subdivision.

There is a special case in the results which may need further explanation. In case of the three times adaptively subdivided Cornell Box with a texture resolution of 64 × 64, rendering with the hemispherical projection is slower than with the non-linear hemisphere projection. As described in Section 2 our subdivision method is based on the ratio between visible and occluded pixels of a triangle. This information is gained from the visibility texture. In this particular example the visibility texture of the hemisphere projection gives incorrect information about shadow boundaries, yielding more subdivisions than actually necessary. In contrast, the non-linear hemispherical projection of Gascuel et al. (Gascuel et al., 2008) gives much better RMS values. The remaining differences are presumable caused by different resolutions used for the hemicube and hemispherical visibility maps and due to the increasing distortions towards the perimeter. It should be pointed out that the non-linear hemispherical projection may suffer from the same sampling problems as the hemispherical projection: small triangles may be not visible in the projected image.

### 5.2 Perceptual Analysis

Although the RMS values quantitatively show the improvement in image quality, they do not reflect if the improvement is sufficient enough to avoid visual differences noticeable by a human observer. Therefore we used the perceptual metric of Yee (Yee, 2004) to assess if there are perceptual differences between the hemicube method and the hemispherical variations. We have chosen this method because an implementation is publicly available. The algorithm requires some user defined parameters for which the standard values were used, except for the field-of-view angle which was set to 37° (assuming an observer looking from 0.5m at a 17” display). An image was considered different if more than 100 pixels (approx. 1.2% of the image) were perceptually different. Since it does not make sense to apply the metric on the lightmaps themselves, we rendered five different views of each scene and applied the metric to the tone-mapped result. Progressive radiosity was turned off because it can change the shooting order and therefore can lead to slightly different results. A solution was considered to have no difference only if all of the views were indistinguishable.

The perceptual metric revealed two problems. First, if silhouette edges are at glancing angles in regard to the projection center, some texels around this edge may be incorrectly classified as in-shadow if one of the two hemispherical projections is used. This seems to be caused by the discretized nature of the visibility map in conjunction with limited depth precision. Sampling the visibility map at the neighboring
texels reduced these problems significantly. Indistinguishable solutions after this modification are highlighted green in Table 2, otherwise they have a red background. The two red entries for the non-linear hemispherical projection are caused by the formerly described problem, that surfaces may not be visible in the visibility map due to distortions. Second, if a hemispherical projection is used, shadows may be smaller or larger than they should be. This was evident in some results because it proved to be inadequate to subdivide only surface elements which are partly in shadow as implemented in our adaption subdivision algorithm. Rather is it necessary to subdivide shadow casting surfaces as well, otherwise their boundaries do not reproduce the circular arc accurately enough (the surface being smaller or larger and therefore allowing more respectively less light to impinge on the receiving surfaces). An example where this problem leads to larger shadows is shown in Figure 5. This issue may be resolved by adaptively tessellating the triangles during creation of the visibility map with a geometry shader. However, this subdivision also has to be considered during back-projection of the individual surfaces. We will look into this issue as part of our future work.

Finally, it should be pointed out that errors in the visibility information may not only affect the image quality of the solution but even the convergence of the radiosity process. In some rare cases we even observed longer rendering times with the hemispheric projection as with the hemicube method, since many areas are incorrectly considered to be in shadow. Figure 6 compares the convergence of a scene, where the errors caused by the stereographic projection increased the number of shootings from 3514 to 4106 and consequently the rendering time from 182.06 seconds to 192.93 seconds.

6 CONCLUSIONS

In this paper we presented an analysis of our GPU radiosity solver. Measurements showed that the performance is mostly independent of the dimensions of the radiosity maps (if the mipmap level for shooting is adjusted accordingly to keep the number of shootings identical). The implementation would benefit if reading from and writing to the same texture location were possible. In most cases the GPU implementation is considerably faster than the CPU implementation. The cost for setting up the required OpenGL re-

\[1\] This mostly occurs near the perimeter of the hemisphere where distortions are very common.

sources relativises with increasing number of radiosity elements and iterations.

Although a hemispherical projection performed in the vertex shader accelerates the creation of the visibility textures, it has several drawbacks which have an impact on the quality of the result. First, the relative position information of overlapping surfaces is distorted, leading to incorrect shadow boundaries. Second, extreme deformations may lead to incorrect visibility classification, misleadingly considering complete surfaces as not visible. Third, if the shadow boundaries are incorrect due to erroneous visibility information caused by the hemispherical projection,
then the subdivision method may unnecessarily subdivide triangles which in turn leads to longer rendering times. A real non-linear projection overcomes almost all of these drawbacks and at the same time decreases rendering times considerably. However, the non-linear projection still exhibits some distortion and sampling problems which are intrinsic to hemispherical projections. Finally, we proposed possible solutions to problems which were revealed during the evaluation and which will be part of our future work.

REFERENCES


