

Sensor-Based Lighting

Sibgrapi paper ID: 114583



Fig. 1. Three indoor-scene renders computed by the Sensor-Based lighting technique

Abstract—Commonly used global illumination techniques have good quality but are computationally expensive and difficult to parallelize in GPU. We propose a novel global illumination algorithm that locates virtual cameras called sensors in a carefully selected set of visible pixels of the scene. Each sensor captures the incoming light from every direction of the scene and produces an estimation value that is later composed with the original image. The algorithm is a highly parallelizable offline solution implemented in GPU, that supports arbitrary-shaped light sources and produces soft shadows, ambient occlusion and color bleeding effects.

Keywords-Radiance estimation; GPU; Global illumination;

I. INTRODUCTION

Many techniques used in computer graphics to simulate lighting effects are based on source lights called point-lights. They are fast to compute and easy to integrate in a real-time graphics pipeline.

However, real life lighting appliances are normally designed to avoid direct lighting because it is uncomfortable for the human eye. For example, a table lamp is usually covered by a surface that produces a glimmering light in the room. Sunlight passing through window curtains behaves like a large superficial light source.

These arbitrary-shaped light sources behave in a very different way from commonly used point lights, for example when generating soft shadows with complex penumbra areas.

Furthermore, in a closed room, the light coming from different light sources will bounce against many surfaces before finally reaching the eye. In each of this bounces part of the energy is absorbed and another part is reflected in certain wavelength, according to the BRDF [1] distribution. The human eye is particularly sensitive to this kind of indirect illumination. This way, a red wall may transfer some of its color to a white ceiling, producing a reddish look, known as color bleeding.

Contributions: The aim of this paper is to present an offline algorithm implemented in GPU that supports all this

subtle effects, from a physically based perspective. This algorithm is capable of modeling a large amount of arbitrary-shaped source lights producing physically correct soft shadows. Moreover it is able to approximate another aspects of global illumination, like color bleeding and ambient occlusion.

II. RELATED WORK

Many algorithms have been developed to generate realistic images based on global illumination models. Path Tracing [2], Monte Carlo Ray Tracing [3] and Metropolis Light Transport [4] are examples that use random sampling to estimate a solution that converges to a physically correct result. But they tend to require thousands of samples per pixel in order to achieve good results, which makes them computationally expensive.

Radiosity [5], is a finite element method (FEM) that models light transport between diffuse surfaces. However, it is also computationally expensive and it requires the scene to be fairly tessellated.

All the previously mentioned techniques aim to produce a physically correct image regardless of the time required to be computed. On the other side, there are faster algorithms that can be used in real-time applications. They can compute direct shadows (shadow map [6], volume shadows [7]) and ambient light (ambient occlusion [8]) in a very simplified form. These techniques are usually known as being perceptually correct. This means that although they are not accurate in physical terms they can be visually attractive to the human eye.

Many methods exist that also produce perceptually correct soft shadows, like Variance Shadow Maps [6] and Cascaded Shadow Maps [7]. Recently, Radiosity and other global illumination algorithms have been adapted to run in GPU, some of them even at real-time frame rates. Instant Radiosity [9] constructs random paths from the light sources and creates new virtual point lights (*VPLs*) where these paths encounter a surface. The work described in [10] uses an hemispherical projection to capture radiance in a path texture.

III. SENSOR-BASED LIGHTING

A. Algorithm overview

The technique presented in this work consists in positioning virtual cameras in selected visible points of the scene, Fig. 2. These cameras act like sensors that measure the amount of incoming light from all directions. In order to achieve that, each camera draws the complete scene in an auxiliary texture. The texture size is carefully chosen in order to capture every illumination detail, and then is reduced to a unique average value which represents the light received from the entire scene. This value is stored in an output texture that is used later to draw the final scene.

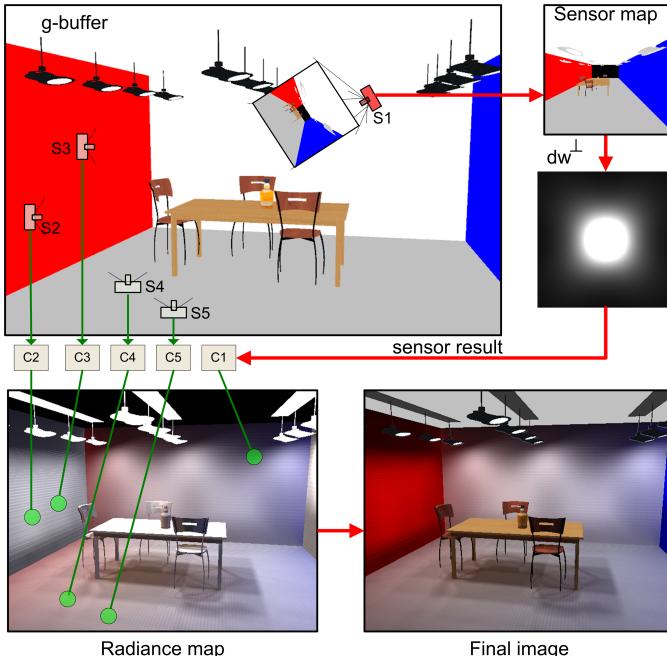


Fig. 2. Sensing function: The light coming from the scene is captured by the top right sensor in the first image. The sensor image is then weighted by its incident angle, and the last step is to average the final color.

B. Radiometry equations

The rendering equation [2] describes the light transport from one surface point to another as the sum of emitted radiance and reflected radiance. Because light transport is usually perceived in equilibrium, the equation describes the steady-state radiance function for the given scene [2].

$$\underbrace{L_o(p, w_o)}_{\text{Outgoing radiance}} = \underbrace{L_e(p, w_o)}_{\text{Emitted Radiance}} + \underbrace{L_s(p, w_o)}_{\text{Scattered Radiance}}$$

The term L_e belongs to the scene description and corresponds to the light sources, which are the only objects that emit light. The second term, the scattering equation, is of more significance for this work:

$$L_s(p, w_o) = \int_{\Omega} L_i(p, w_i) f_s(p, w_i \rightarrow w_o) d\omega^{\perp}$$

Where w_o is the view direction, w_i the incoming direction, $d\omega^{\perp}$ is the differential projected solid angle and Ω represents the positive hemisphere. The f_s term is the BRDF [1] and represents the amount of radiance coming from the w_i direction that reflects towards the w_o direction. In its simplest case the function can be constant (perfect diffuse surfaces) and the f_s becomes a $\frac{c}{\pi}$ factor out of the integral [11, p. 78]. c is a constant ≤ 1 and shows the fact that the radiance reflected out of the surface can not be greater than the incoming radiance.

The ingoing radiance can be expressed in terms of outgoing radiance using a ray casting function [11, p. 90], so the scattering equation becomes:

$$L_s(p, w_o) = \frac{c}{\pi} \int_{\Omega} L_o(ray(p, w), -w) d\omega^{\perp}$$

This equation can be used to predict the surface color from the incident light [11, p. 86]. Our work applies this principle to adapt the equation to a GPU suitable computation, Fig. 3.

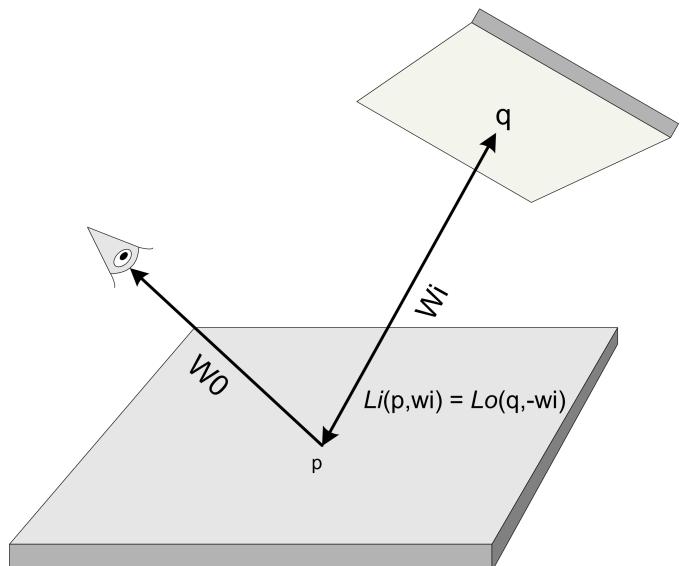


Fig. 3. Light Transport. Ingoing Radiance (L_i) can be expressed in terms of Outgoing Radiance (L_o). The radiance entering to p in w_i direction is the radiance leaving from q with $-w_i$ direction. q is the first intersection between the ray from p towards w_i direction and the rest of scene.

C. Outgoing radiance estimation

L_o appears in both sides of the rendering equation, forming a recursive expression. The strategy of this work is to replace L_o by an estimator to avoid the recursive form. Considering a fixed direction w_i and the point $q = ray(p, w_i)$ then two situations must be studied:

A If q is in the surface of a light source, L_e (emitted radiance) will have a value corresponding to the color emitted by that light, which belongs to the scene. We can assume that this value is considerably higher than L_s for a light source. Thus:

$$L_o = L_e \quad (1)$$

B If q is in a surface that reflects light, L_o is described by the scattering equation, then its value depends on the light coming from all directions of the hemisphere. In this case the surface diffuse color (also part of the scene description) is taken as an estimation of the value, multiplied by an epsilon e:

$$L_o = D \cdot e \quad (2)$$

The epsilon term is a small number that allows us to set the importance of the ambient light related to the direct light. It also represents the fact that the incident light is usually reflected in many directions, so the amount of light reflected in one particular direction is very small [1]. So the L_o estimator becomes:

$$L_o^* = \text{Color}(q)K(q)$$

The estimator combines 1 and 2 in the same equation. From a computational point of view the difference between 1 and 2 is just the specific value of $K(p)$. In one case it is the light associated energy and in the other it is a small epsilon. Nevertheless, from a physical point of view both cases correspond to entirely different processes. It is important to note that the calculated values may take a wide range of results, so the precision of a high dynamic range [12] is required. The term $\text{Color}(q)K(q)$ will be shortened as $Ck(q)$.

D. Scattered radiance estimation

Replacing L_o^* in the scattering equation and approximating the integral with a summation we get 3:

$$L_s(p, w_0) = \frac{c}{\pi} \int_{\Omega} Ck(q)d\omega^{\perp} \sim \frac{c}{\pi} \sum_{i=1}^n Ck(q_i)d\omega^{\perp} \quad (3)$$

E. Sensor

In this algorithm a sensor represents a virtual camera that allows us to capture the incoming light from all directions of the scene. This light is then stored in a texture named *sensor map* that represents the sensor area. The camera is located at the required point p in the direction normal to the surface of the scene, Fig. 4.

The *sensor map* size is $rd_x \times rd_y$ and each texel represents a sample of the Ck term for a given direction in space. Hence, the sensor is applied to compute the Ck term for a large quantity of samples and then to use them in the equation 3. In order to be able to capture the scene illumination from as many directions as possible, the camera field of view (*fov*) is set to 135 degrees, instead of the commonly used 45 degrees. This brings out two main problems. Firstly, such a high value in *fov* amplitude results in a loss of precision at the center of the image, just where the samples have more incidence. Secondly, we can never take a sample of 180 degrees of amplitude using a perspective or orthogonal projection. This means the sensor is incapable of capturing the light coming from angles out of the *fov*. The Ck term for a giving pixel, drawn from the sensor point of view, can be computed using a Pixel Shader in GPU because both *Color* and *K* are part of the scene description. The original summation over the amount

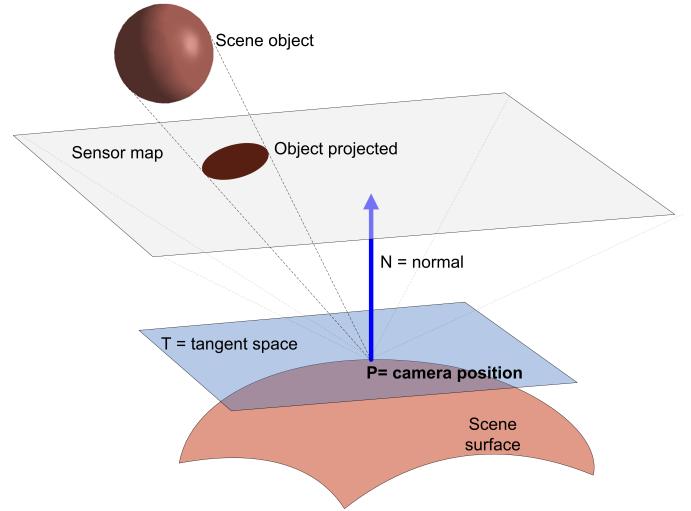


Fig. 4. Sensor setup. The virtual camera is located at the required point p in the direction normal to the scene surface.

of samples is replaced by a double summation to iterate over the values of the *sensor map* in width and height.

The remaining part of the equation, dw^{\perp} (differential projected solid angle), can be obtained by simply projecting the area of the unit sphere covered by the dw (differential solid angle) orthogonally onto the tangent space, and then finding the area of the resulting planar region (Nusselt's Analog [11, p. 78]).

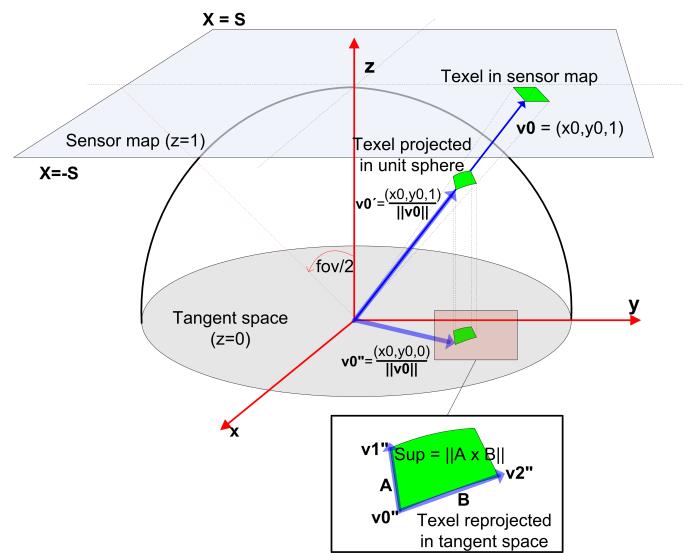


Fig. 5. Solid angle Solid angle measure and projected pixel area. A texel area in the *sensor map* is first projected in the unit sphere and then projected again in tangent space. This projected area is used to measure the contribution of the sample.

As we can see in image Fig. 5:

$$S = \tan\left(\frac{\text{fov}}{2}\right)$$

$$x = 2S\left(\frac{j}{rd_x} - 0.5\right)$$

$$y = 2S\left(\frac{i}{rd_y} - 0.5\right)$$

An arbitrary point $v = (x, y, 1)$ in the *sensor map* is first normalized to project it in the unit sphere, and then is projected again in tangent space by simply setting $z = 0$. We denote this with the " operator:

$$V'' = \frac{(x, y, 0)}{\|(x, y, 1)\|}$$

Given a sample (i,j) in the sensor area, we call V_0 , V_1 and V_2 to the vertex:

$$V_0 = (x_j, y_i, 1), V_1 = (x_{j+1}, y_i, 1), V_2 = (x_j, y_{i+1}, 1)$$

The projected pixel area in tangent space can be approximated using de cross product:

$$d = \|(V_1'' - V_0'') \times (V_2'' - V_0'')\| \quad (4)$$

Equation 4 represents the pixel projection over the tangent space in arbitrary pixel units. The sum of the area of all possible angles is equal to the whole disc surface hemisphere, which equals π . So the relative value of each sample (without units) is:

$$\frac{\|(V_1'' - V_0'') \times (V_2'' - V_0'')\|}{\pi}$$

The total area of the samples can be calculated as $rd_x \cdot rd_y$, which allows us to approximate the term (4) with:

$$F_{ij} = \frac{\|(V_1'' - V_0'') \times (V_2'' - V_0'')\|}{\pi} \cdot (rd_x \cdot rd_y)$$

With this formulation, the energy values associated with each source light must be set in [radiance units] \times pixel.

As a result the F_{ij} values only depend on i and j , so they can be pre-computed. The physical interpretation means that the light coming from similar directions to the normal has the max incidence over the sensed value, while the peripheral directions decrease rapidly.

The estimated scattering equation becomes:

$$L_s(p, w_o) \sim \sum_{j=1}^{rd_x} \sum_{i=1}^{rd_y} c_{ij} F_{ij}$$

This value is the one obtained as a result in each sensor and then stored in the output texture called *radiance map*.

F. Visible points determination

The first step consists of creating a geometry buffer to store some useful information about the visible points of the scene: position, normal and polygon or mesh id. Different packing strategies derived from the deferred rendering techniques [13] can be used to create this *G-buffer*.

G. Sensor quantity and distribution

In many parts of the domain the value of L_s changes in a soft way. Except when there is a change of surfaces, it is very likely that L_s will be continuous in the point neighborhood. This suggests that it is not necessary to compute L_s in all the image points. The algorithm can sense at regular intervals from a given tile of $dt \times dt$. Each tile will have a sample to which L_s is estimated. Then for each sample it computes the variation with the four closer samples, Fig. 6. If this variation is less than a given threshold, called *dradiance*, the whole set can be estimated with an interpolation between the four neighbors. On the other hand, if the variation is greater than *dradiance* then more samples may be located in tiles of half the size of the initial interval ($\frac{dt}{2}$). The process is repeated until only final pixels are left, where L_s is computed specifically to complete the scene, Fig. 7. In this way the algorithm converges to the solution in a progressive way, using more sensors where the image needs them most.

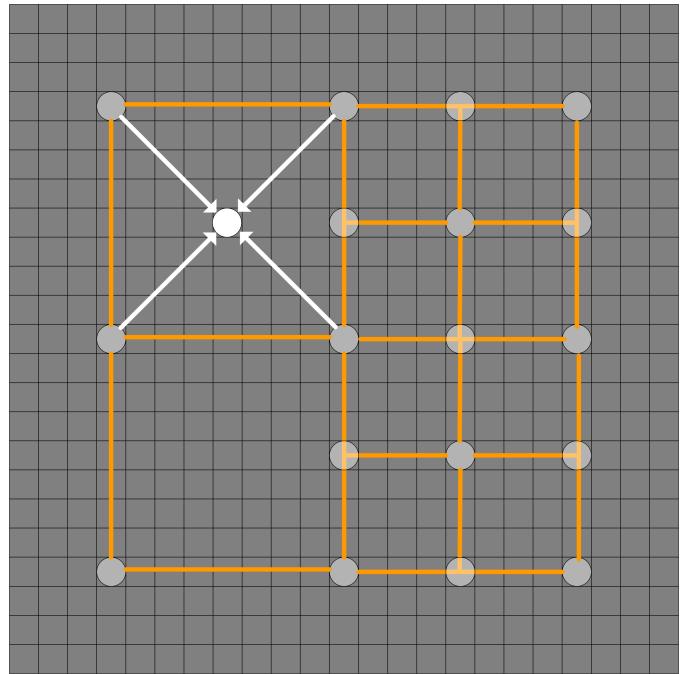


Fig. 6. Bilinear interpolation of the underlying pixels from the calculated values in the quad. The samples with little variation can be interpolated to drastically reduce the amount of sensing operations.

H. Final image composition

So far we have generated a map of values for L_s for each visible point of the scene, which are stored in the *radiance map*. Each value is called $R(x, y)$. Next, the texture is combined with the original image in order to produce the final image. For a given point (x, y) of the original image, the same point of the *radiance map* contains the value of L_s . Both values are multiplied:

$$final(x, y) = color(x, y)R(x, y)$$

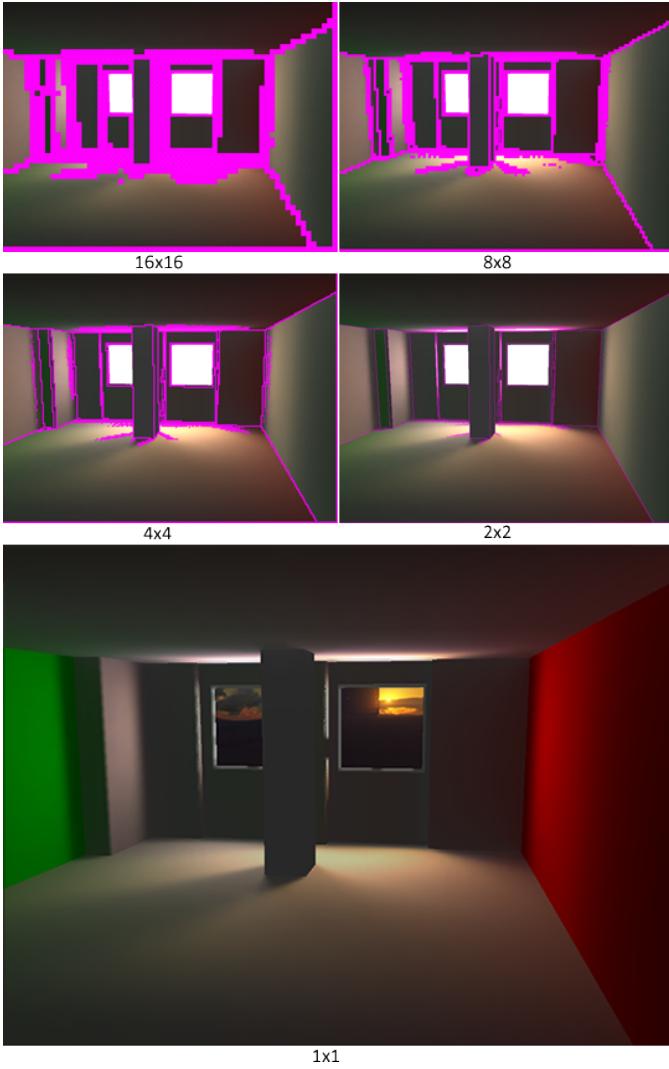


Fig. 7. Tile progressive computation. The first image shows tiles of 16×16 pixels. Only one sample per tile is computed which value is interpolated between the neighborhood samples. In the next step the tile size is reduced. At the end only the edge pixels remain to be sensed, where L_s is discontinuous.

Where $color(x, y)$ is the original diffuse color of the object without any lighting calculation. The tile optimization of the algorithm tends to produce small discrete blocks in the final image, an aliasing normally seen in most sampling techniques that use regular intervals. Replacing the regular sampling by a random one (for example with jittering techniques) will transform those aliasing in noise, but the visible artifacts will be appreciable. A better solution is to apply a Gaussian filter [14] to image composition. Fig. 8. A certain radius of samples of the *radiance map* is taken, provided that they belong to the same surface. This restriction is vital in order to avoid summing values of L_s coming from different regions of space, which will produce hard artifacts in the solution. That is why

the *G-Buffer* stores the polygon *id*.

$$final(x, y) = color(x, y) \left[\sum_{i=-r}^r \sum_{j=-r}^r k_{ij} L_s(x + i, y + j) \right]$$

k_{ij} is the relative weight of each sample and is computed considering the total amount of samples that satisfied the surface *id* condition (that is they belong to the same polygon as the point (x, y)).

$$k_{ij} = \begin{cases} \gamma_{ij} & \text{if } id(x + i, y + i) = id(x, y) \\ 0 & \text{if } id(x + i, y + i) \neq id(x, y) \end{cases}$$

$$\sum_{i,j=-r}^r \gamma_{ij} = 1$$

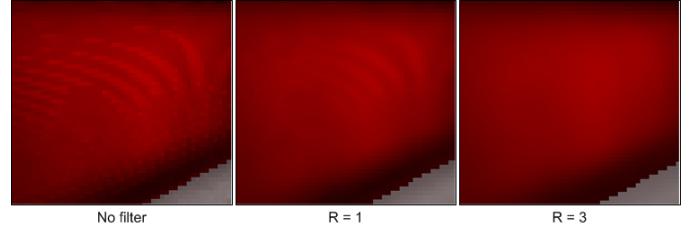


Fig. 8. Gaussian filter application in image composition. Note that the red surface edge (wall) and the gray one (floor) continue to be sharp. The filter is specifically designed to sky areas where the *id* changes suddenly.

I. Shadow computing

The size of the *sensor map* ($rd_x \times rd_y$) must be large enough to capture all the details from the light sources. An area too small may generate discontinuities in the estimation of L_s which will be perceived as noise in the image.

For superficial light sources, or for large enough arbitrary-shaped lights, this estimation can be extremely accurate when the proper texture size is given. This is due to the fact that the projection of the light source over the sensor area occupies a large amount of samples. So the sensor is very sensitive to obstacles that may block part of it, producing physically correct shadow penumbras, Fig. 9.

IV. IMPLEMENTATION AND RESULTS

The algorithm was implemented using C++ with DirectX 9 and Shader Model 3. Three different metrics were taken in order to understand the behavior of the technique. The total time spent by the algorithm depends on many factors (screen resolution, *sensor map* size, geometry complexity of the scene), so a better indicator is what we called sensor time, which is defined as the total time divided by the amount of sensors required for that scene.

The first metric, Table I, measures different values for sensor time as the screen resolution increases and sensor map size remains constant. The gathered data shows that the time spent BY each sensor is constant for a fixed *sensor map* size, and it does not depend on the screen resolution.

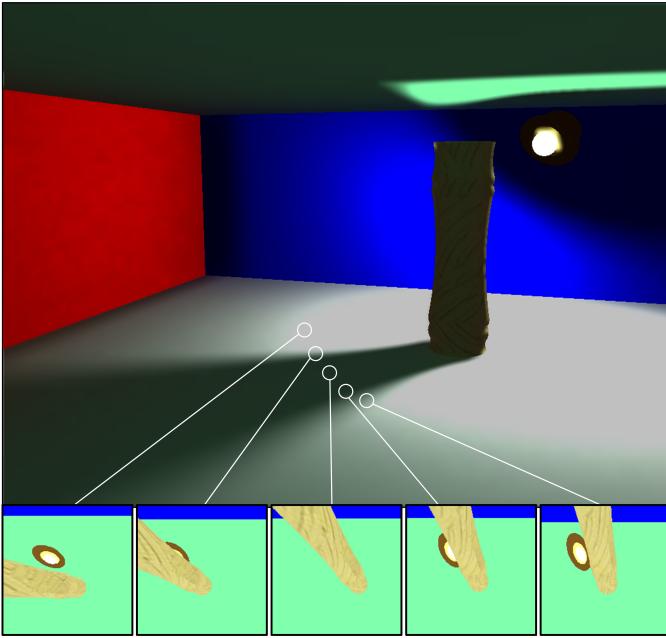


Fig. 9. Shadow computing. The circle represents the sensed area. The first screenshot shows the light source totally visible. Then, the visible area of the same gradually decreases as the shadow becomes more evident in the final image. In the third screenshot the light source is no longer visible and the area is completely in shadows.

TABLE I
FIRST METRIC: SENSOR TIME RELATED TO SCREEN RESOLUTION.

Screen resolution	Sensors count	Time/sensor (ms)
200x200	15,240	4.455
300x200	23,762	4.439
400x300	40,271	4.499
500x300	44,300	4.496
500x400	55,904	4.541
600x400	63,074	4.581
600x500	68,276	4.531
700x500	79,158	4.587
700x600	83,423	4.588
800x600	96,266	4.588
900x600	108,239	4.578
900x700	112,436	4.581
1024x768	133,822	4.569

The second metric, Fig. 10, measures the sensor time as the *sensor map* size is increased. Three different screen resolutions were taken as reference. The conclusion is that the *sensor map* size influences directly over the total processing time of the algorithm, almost in a quadratic fashion, and it does not depend on the screen resolution.

The third metric, Fig. 11, measures the amount of sensors required by the scene as the *sensor map* size is increased. The screen resolutions used in this test were the same as the ones used in metric two. The amount of sensors is high when the sensor map is small, because each sensor has not enough resolution to capture the scene details, so more sensors are required. But after increasing the size to 128×128 , the sensors count begins to stabilize. In our tests, we achieved a good compromise between quality and performance with a *sensor*

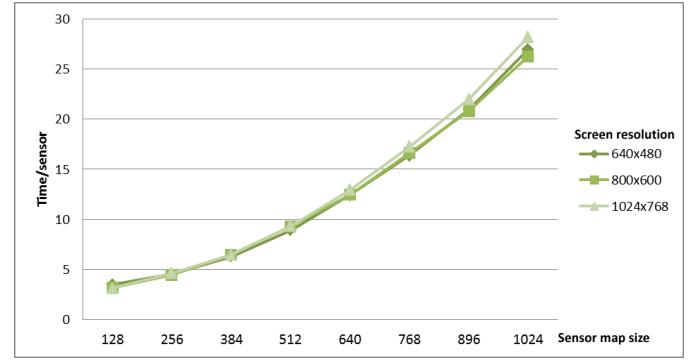


Fig. 10. Second metric: sensor time related to the *sensor map* size

map of 256×256 , Fig. 12. The tests performed also show that the total time required by the algorithm is much more sensitive to the amount of sensors than to the geometry complexity of the scene.

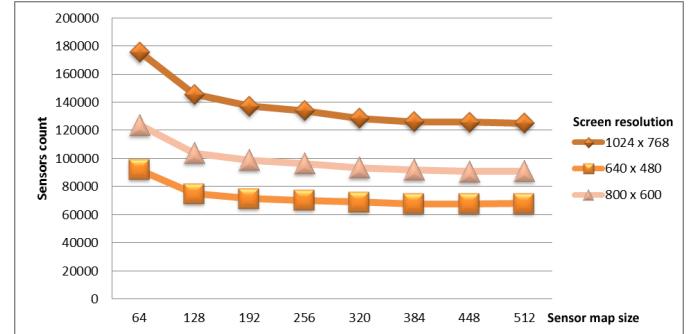


Fig. 11. Third metric: sensors count related to *sensor map* size.

The results were computed using a PC with Intel Core i7 3.50GHz processor with 16GB RAM and NVIDIA GFORCE GTX 660 GPU, with Windows 7 64 bits.

V. CONCLUSIONS

The algorithm presented provides an alternate method to solve the global illumination of the scene, with support for complex effects like arbitrary-shaped light sources, soft shadows, ambient occlusion and color bleeding. The solution is designed to be computed in GPU, taking advantage of many hardware accelerated features. Only one pass is required and it does not need to store any intermediate value, which makes it a highly parallelizable solution. However, a minimal synchronization is required at the end of each tile pass in order to interpolate the values.

The final image obtained by the method has a very realistic looking and it requires relatively less processing time than many Ray-Tracing based solutions [15]. Unlike Radiosity, it is not required to heavily tessellate the scene, which represents one of the drawbacks of that technique. The original scene data can directly be used by the algorithm.

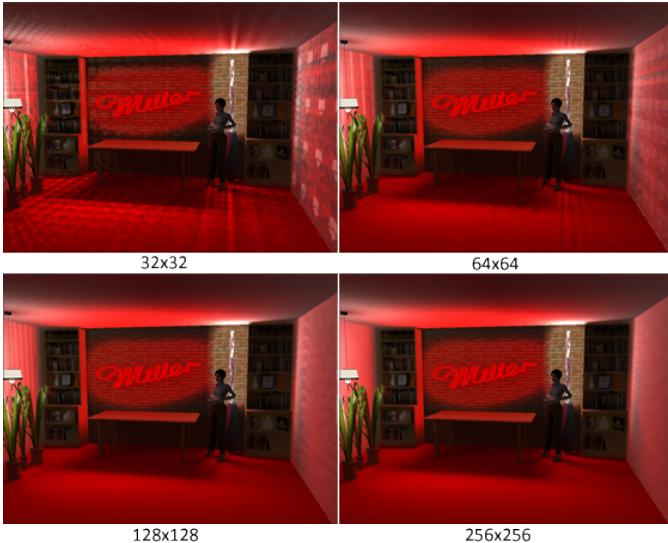


Fig. 12. The sensor area. Left: a 32×32 pixel sensor. Note the noise on the floor and the left wall.

VI. FUTURE WORK

The technique is in its early stages. Further studies for optimization are required in order to achieve the performance and scalability of professional rendering solutions.

Its main problems are produced by the limited *fov* used by the sensor, which makes the algorithm incapable of capturing the illumination from certain angles. Also, point light sources may be difficult to compute since they have to be modeled as a very small surface with a high energy level, making them hard to capture by the sensor. Both problems could be solved by using ray tracing pointed to the light sources to estimate a statistical value of visibility degree for that light source in the problematic angles. Few rays should be needed so the performance of the solution should not be overly affected.

Some objects of the scene are not good candidates to shade with a global illumination technique like the one presented in this work. For example very irregular surfaces, a human face, highly specular surfaces. This type of objects could be excluded of the algorithm and computed with a better suitable technique, substantially reducing the amount of sensors required. Our illumination technique allows this kind of flexibility because each pixel is completely independent of the rest of the scene.

Some others forms of BRDF could be easily incorporated in the solution replacing the constant factor by a mixed equation. This equation may contemplate diffuse surfaces and some partially glossy mirrors surface. Perfect mirrors would be problematic because the variance of the solution would produce too much noise in the image. Others techniques suit better for this kind of surfaces.

REFERENCES

- [1] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis, "Radiometry," L. B. Wolff, S. A. Shafer, and G. Healey, Eds. USA: Jones and Bartlett Publishers, Inc., 1992, ch. Geometrical considerations and nomenclature for reflectance, pp. 94–145. [Online]. Available: <http://dl.acm.org/citation.cfm?id=136913.136929>
- [2] J. T. Kajiya, "The rendering equation," in *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '86. New York, NY, USA: ACM, 1986, pp. 143–150. [Online]. Available: <http://doi.acm.org/10.1145/15922.15902>
- [3] K. Suffern, *Ray Tracing from the Ground Up*. Natick, MA, USA: A. K. Peters, Ltd., 2007.
- [4] E. Veach and L. J. Guibas, "Metropolis light transport," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '97. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 65–76. [Online]. Available: <http://dx.doi.org/10.1145/258734.258775>
- [5] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile, "Modeling the interaction of light between diffuse surfaces," in *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '84. New York, NY, USA: ACM, 1984, pp. 213–222. [Online]. Available: <http://doi.acm.org/10.1145/800031.808601>
- [6] L. Williams, "Casting curved shadows on curved surfaces," in *Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '78. New York, NY, USA: ACM, 1978, pp. 270–274. [Online]. Available: <http://doi.acm.org/10.1145/800248.807402>
- [7] E. Eisemann, M. Schwarz, U. Assarsson, and M. Wimmer, *Real-Time Shadows*. A.K. Peters, 2011. [Online]. Available: <http://www.cg.tuwien.ac.at/research/publications/2011/EISEMANN-2011-RTS/>
- [8] T. Akenine-Moller, T. Moller, and E. Haines, *Real-Time Rendering*, 2nd ed. Natick, MA, USA: A. K. Peters, Ltd., 2002.
- [9] A. Keller, "Instant radiosity," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '97. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 49–56. [Online]. Available: <http://dx.doi.org/10.1145/258734.258769>
- [10] M. Pharr and R. Fernando, *GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation (Gpu Gems)*. Addison-Wesley Professional, 2005.
- [11] E. Veach, "Robust monte carlo methods for light transport simulation," Ph.D. dissertation, Stanford, CA, USA, 1998, aAI9837162.
- [12] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.
- [13] G. Liktor and C. Dachsbaecher, "Decoupled deferred shading for hardware rasterization," in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, ser. I3D '12. New York, NY, USA: ACM, 2012, pp. 143–150. [Online]. Available: <http://doi.acm.org/10.1145/2159616.2159640>
- [14] S. Molnar, "Efficient supersampling antialiasing for high-performance architectures," DTIC Document, Tech. Rep., 1991.
- [15] P. S. Heckbert, "Graphics gems iv," P. S. Heckbert, Ed. San Diego, CA, USA: Academic Press Professional, Inc., 1994, ch. A minimal ray tracer, pp. 375–381. [Online]. Available: <http://dl.acm.org/citation.cfm?id=180895.180928>

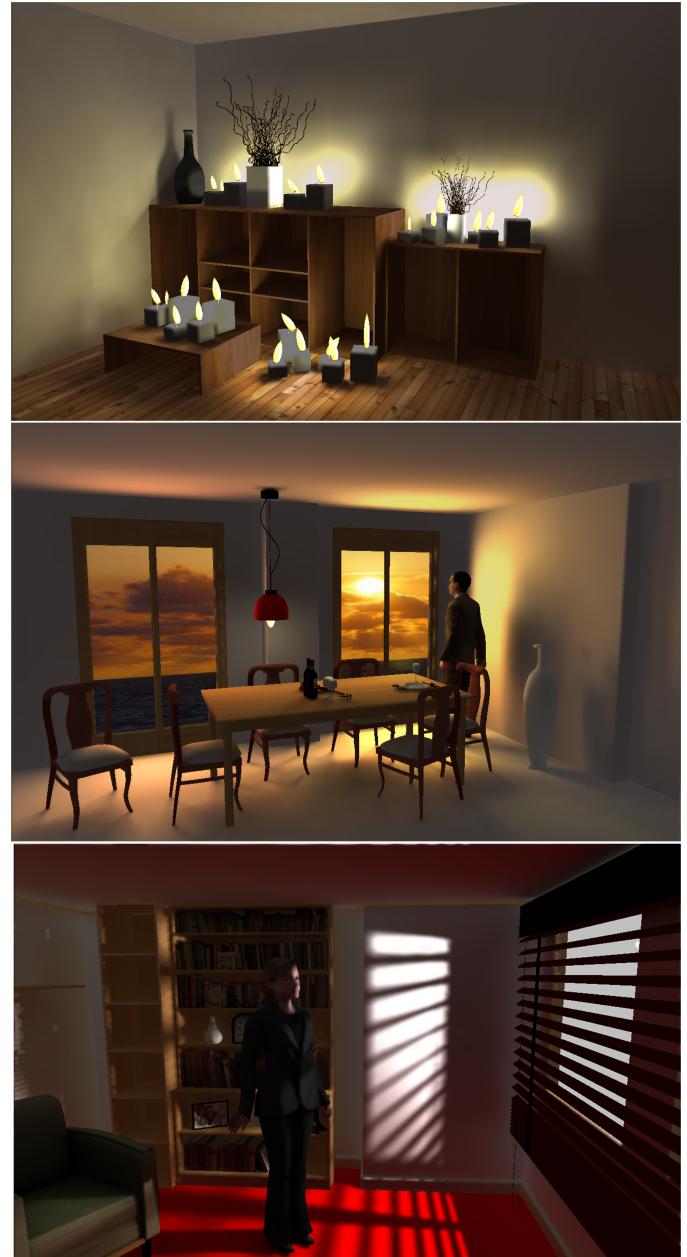


Fig. 13. Screenshot taken by the Sensor-Based Lighting technique

Fig. 14. Screenshots taken by the Sensor-Based Lighting technique