A Hybrid Ambient Occlusion Technique for Dynamic Scenes

Martin Knuth
Fraunhofer IGD
Fraunhoferstr. 5
64283 Darmstadt, Germany
martin.knuth@
igd.fraunhofer.de

Jörn Kohlhammer
TU Darmstadt
Fraunhoferstr. 5
64283 Darmstadt, Germany
Joern.Kohlhammer@
gris.informatik.tu-darmstadt.de

ABSTRACT

In this paper we present a hybrid technique for illuminating a dynamic scene with self shadowing. Our goal is to perform the necessary calculations with real-time or interactive frame rates. The main idea is to split up the self shadowing process into a global and a local part. This separation allows us to choose combinations of completely independent algorithmic approaches. The global part calculates the self shadowing information of the entire scene. Since it has to process the whole scene, it has a high computational cost; however, a coarse approximation can be created in a short time. In contrast the process for the local part deals with fine details, which have local impact on the scene. It only processes a relevant subset of the scene. Finally we present results based on an implementation using a GPU based self shadowing approximation combined with screen space ambient occlusion for solving the local part

Keywords Shadow Algorithms, GPU Programming, Ambient Occlusion

1. INTRODUCTION

Shadows are a very important part for the realism of a 3D scene. They give information about spatial object relations, allowing better immersion. Our goal is to illuminate a fully dynamic 3D scene with a high dynamic range (HDR) light probe surrounding this scene. Since the light probe describes colour and intensity for given directions, it is necessary to know the directional occlusion of the light for each position of the scene. Additionally the lighting condition described by a light probe creates a wide range of shadow frequencies. This is especially a problem for dynamic and deformable scenes, where no information about self occlusion can be precomputed. To cope with the complexity of the shadowing problem in large, dynamic scenes, we propose a new approach. The core of the problem is based on the fact, that all geometry and all lighting parameters may affect the lighting conditions in every single point in the scene. All approaches are approximations, working on specific samples of geometry, spatial areas or light sources (or combinations thereof), always resulting in a trade-off between computational speed and accuracy/realism. The strategy for our approach is to use two shadowing algorithms for two different computational domains: A global domain for the dependencies between the lighting conditions of the complete scene and a local domain for the small-scale dependencies especially focussing on the visible part of the scene. In the global domain the light exchange is approximated on a scale coarse enough to allow real-time rendering. The computational effort for the details of the scene is shifted to the algorithm operating on the local domain. The algorithms work and can be chosen independently from each other. The result of both algorithms is a ”shadow-mask” of the scene, defining the effect for its specific domain. We combine the two results, in the final step.

In our implementation section we present the implementation of a particular combination of algorithms. This combination shows self shadowing evaluation per pixel and the ability to trade speed against accuracy in the global part. In the local part we decided to use a screen space ambient occlusion (SSAO) technique, which has the advantage of being independent from the complexity of the scene.

The presented implementation processes the all frequency self shadowing of the scene at interactive or real time frame rates on current GPUs. It is able to process non-manifold deformable objects within fully dynamic scenes without prior knowledge of the animation.

In section 2 we refer to work related to our approach. Additionally we refer to two previus works, that we use for the global and the local computation. Then, we explain our technique and considerations behind it. In section 5 we present the particular case we have chosen and implemented. In the next section we present a summary and results of this approach, followed by a future work section.

2. RELATED WORK

We will now present a collection of algorithms, which aim at similar goals. These algorithms can be classified as follows. We will start with algorithms, which deal with the shadowing problem of single light sources. A survey of real time shadowing methods is given in [HLHS03]. Some of the presented algorithms work with small area light sources, but can be computational costly if these lights become large. Fast shadow calculations for simple point or directional lights within dynamic scenes can be done inside modern GPU's. Mainly two different approaches are used: Shadow volumes [Cro77] and shadow maps [Wil78]. Shadow maps are fast to compute and need less fill rate than shadow volumes. On the other hand, care has to be taken of sampling artefacts. An approach for minimizing shadow buffer artefacts and softening of the borders of shadows can be found in [RSC87, DL06].

We will now refer to algorithms, which provide or store directional lighting information. If shadows are not necessary, the scene can be directly illuminated by an environment map. For this task a lot of fast environment mapping methods (like [Gre86, RH01, HS99])exist. The advantage of using environment map based approaches is based on their capability to consider all lights at once, which makes them independent from the number lights used. High dynamic range (HDR) environment maps [Deb98] represent the full variablility of natural lighting conditions, which can be used to bring together synthetic and real scenes. An environment map used for illuminating a scene is referred as lightprobe in this literature. Since this technique does not support directional shadowing directly it can be necessary to compute a carefully chosen set of light sources to approximate the lighting effect of a light probe. As an alternative methods can be used, that handle directional information directly. The represented occlusion or lighting data is then used to process the environment map. These methods store directional data, using spherical harmonics, or similar methods. The type of data used often require precomputations, so they cannot be used for animated geometry. This problem is addressed by Kautz et al. in [KLA04]. By using model hierarchies they accelerated the calculations to interactive frame rates. Limitations are small model sizes and low frequency shadows. In addition model hierarchies are costly to compute. A method with similar results is presented in [SSZK04]. It evaluates the visibility of a given set of directional light sources per vertex, taking into account the colour of the environment map, by sampling the map in an uniform way. Visibility calculation is done inside the GPU, and is read back to the CPU which is a bottleneck.

At last we will refer to ray casting and ray tracing methods, which too are well suited to answer occlusion questions. [WPS$^+$03, WBS03, PBMH02]. Nowadays, real-time frame rates can be achieved by using a PC-cluster, the GPU or dedicated ray tracing hardware, which is still in development. Software solutions require the power of an efficient cluster system and at last GPU ray tracing needs to solve the bottleneck problem of reading back towards the CPU.

Utilising a real time ray tracer, global illumination can be simulated. In [CHL04] a direct approach towards global illumination on the GPU is presented, which works for small scenes. Another approach aimed for fast global illumination computations was presented by Keller et. al in [Kel97]. In this method single shadow casting lights are used to illuminate the scene.

Previous Work

We will now describe two techniques which we used as building blocks for the local and the global part in our implementation.

In [KF05] a fast GPU based method for evaluating the directional occlusion of a 3D scene is described. The method heavily relies on the capability of modern GPU to quickly perform shadow calculations. The HDR light probe is approximated by a set of light sources, which are used in conjunction with shadow mapping. The advantage of this approach lies in the possibility to tune its performance with the parameters "'light sources"' and "'shadow map resolution"'. Since the map resolution is directly related with the details preserved within the casted shadows this is a complement to the SSAO algorithm. However one drawback of the presented method is an effect called "'light bleeding"' which occurs, when using shadow

 ISBN 978-80-86943-94-7

maps with low resolutions. To prevent aliasing artefacts the masking effect of the shadow buffer is applied a short way behind the occluder to not accidently shadow the occluder's front.

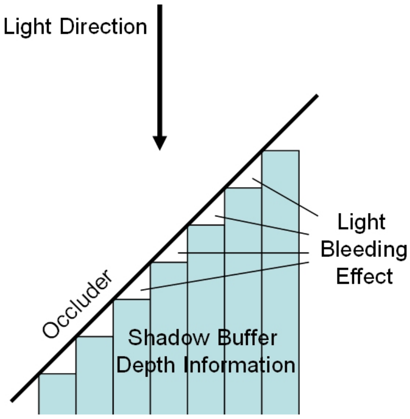

Figure 1: To prevent shadow texel aliasing effects the shadow effect is moved a bit behind the occluder. This results in a "'light bleeding"' effect for thin occluders, since the light travels a bit through the occluder, until it is masked out by the shadow buffer.

To deal with the local occlusion we use a SSAO technique. The key idea of (SSAO) is to interpret the depth information as a height field. For this height field the self occlusion is computed. The approach is decoupled from the scene geometry, since it is completely screen based. In [SA07] the authors present a technique for evaluating the ambient occlusion (AO) of a scene. They divide the ambient occlusion problem into a near AO part and a far AO part (these stand for the high and low frequency details of the scene). For the low frequency part they use an arrangement of spheres which allows a quick approximation of the ambient occlusion in a very coarse way. The authors evaluate the high frequency part by a SSAO technique. This algorithm uses the depth information of the current frame. However the depth information of the current scene is only a very lossy representation of the information defined by the 3D scene. The localness of the method is given by the range in which occluders are searched around the currently processed object.

In difference to this technique our presented combination is able to use a HDR lightprobe to model the incoming light. Ambient Occlusion assumes a steady over the area of the lightprobe. In combination with lightprobes representing an unbalanced intensity distribution this can lead to a wrong illumination.(figure 2)

At last our implementation makes massive use of deferred shading in order to separate the computations from the scene complexity. The idea of deferred shading is originally presented in [DWS+88]. The key idea behind deferred shading is to store rendering information in images once per frame. These images resample information of the rendered objects, like depth, normal, colour and other scene data. This allows to perform processes like "'lighting the scene"' without the need to touch the original geometry. Typically these images are rendered into a so called multiple render target (MRT), which allows the rendering of several images at once. Algorithms, which only need to access this data, run independently of the scene complexity.

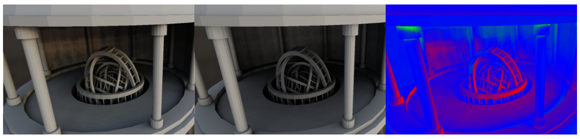

Figure 2: Difference between the occlusion of single light sources and ambient occlusion. Left: shadowing of the single light sources. Middle: Ambient occlusion. Left: difference image. The beach-cross lightprobe[Deb98] was approximated by 2048 light sources to create a high quality rendering

3. HYBRID OCCLUSION

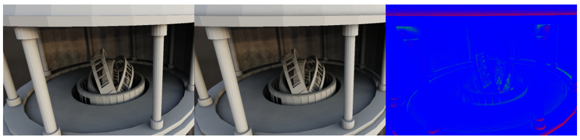

Figure 3: Difference between the occlusion of single light sources and an ideal hybrid directional/ambient occlusion combination. Left: shadowing of the single light sources. Middle: Result of an ideal hybrid rendering. Left: difference image. Shadows of the single light sources are computed on a coarse level to reduce computational costs. Details are processed by a local Ambient Occlusion technique. In Difference to 2 the difference image shows fewer artefacts, since Ambient Occlusion now only affects small regions.

Our key idea is to separate the complex self shadowing problem into two parts. The first step is the global shadow calculation on a level coarse enough to allow interactive frame-rates. We use a shadow map based approximation for the directional occlusion, but our hybrid approach is not restricted to this specific class of algorithms.

The benefit of this approach is mainly given by the characteristics of the local part. In the local part occluders are only tested near the current object of interest. This reveals high frequency details only and has the low computational costs of a local process. The

 ISBN 978-80-86943-94-7

main strategy is the reduction of overall computational costs by reducing the costs in the global part far more then the costs appended by the local part.

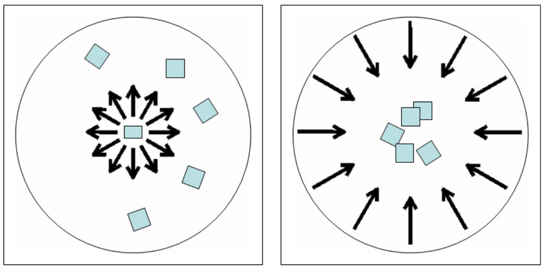

Figure 4: Differences between Local(right) and Global(left) Part. In the local part an object tries to figure out its occlusion. In the global part light from outside the scene is occluded by the scene.

The Global Part

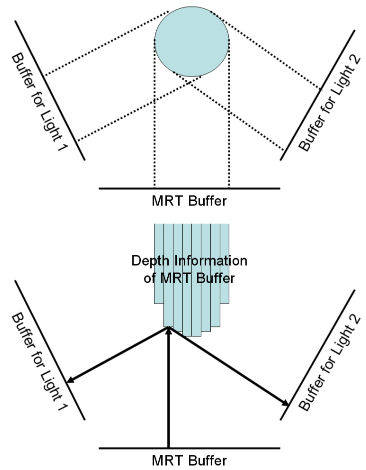

Figure 5: Principle functionality of our global part workflow. Upper image: We first render the whole scene in the scene/shadow buffers. Lower image: in the lighting stage only the data inside the buffers is needed to perform the shadow computations. The world position is calculated from depth and camera information. This position is then transformed into shadow map world space for the lookup.

Our intention within the global part is to calculate an approximation of the directional occlusion effects of the entire scene. The process starts by converting the light probe into a set of directional light sources, for which shadow maps are computed to approximate the directional occlusion. The conversion process has to be done once; since the light probe has a static property. However, the shadowing process itself is dynamic and can be performed completely inside the GPU. Additionally we intent to overcome the appearance of single shadow borders by by adding noise to the sample position of the shadow map. Unfortunately this increases the amount of light bleeding, since the higher threshold has to be chosen in order to not shadow the occluder accidently. This is due to virtually increasing the area of a texel in the shadow map, which introduces the same artefact like decreasing the shadow map resolution. The noising allows us to reduce the shadow map resolution, reducing the fill rate used to generate the shadow maps, saving computational time. Unfortunately the whole shadowing process will tend towards low frequency shadows, since the high frequency parts are blurred away, too. These two issues, light bleeding and missing details are then addressed in the local part.

The Local Part

The idea of the local part is to add missing details to the scene shadowing, and to correct artefacts, introduced by the global parts approximation. In this part we want to retrieve missing high frequency shadows and additionally we reduce the light bleeding artefacts. The basic approach is to apply an algorithm to the output image of the global part, which darkens concave areas of the scene with a local operation. Since this wanted property is similar to ambient occlusion it fulfils both purposes.

Combination

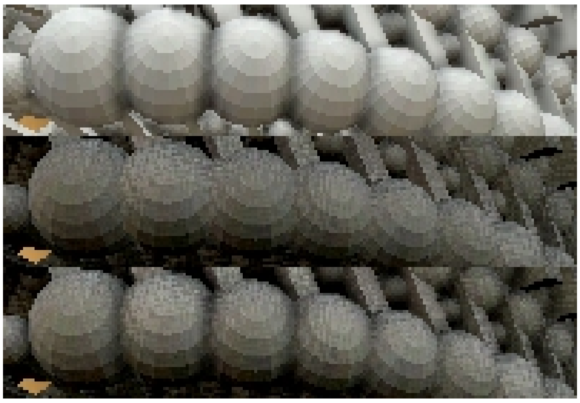

Figure 6: The effects of the details introduced by the local part: Top: local output. Middel: global part output. Down: the final combination.

In a final stage we have to combine the results of the two used methods. To simplify this we constructed the whole process around a deferred shading rendering system. The local part outputs an attenuation value, with which the result from the global pass can be modulated. Modulation is a sufficient operation in this case, since directional information provided by the global part can not be applied to the local part

 ISBN 978-80-86943-94-7

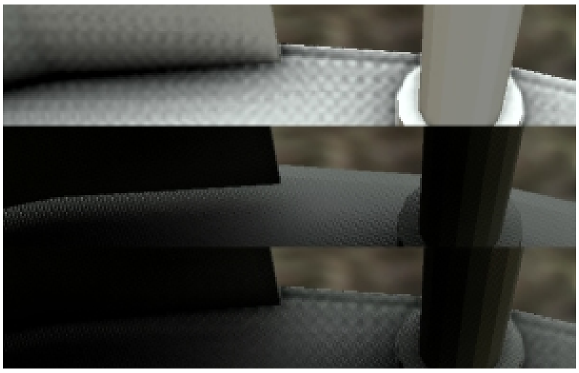

Figure 7: Light bleeding reduction introduced by the local part: Top: local output. Middel: global part output. Down: the final combination.

algorithm. So both outputs are combined in a final stage and not on a per lightsource level. Otherwise the output factors of both algorithms would have to be merged regarding their direction. By using an ambient occlusion technique in the local part we can work around this problem.

4. IMPLEMENTATION

We implemented our technique on a NVIDIA 8k Series graphics hardware. It is a recent Graphics processor, which is capable to use the same processing stages for vertex processing as well as fragment processing. We use this capability to perform most of the work on a per fragment level.

Global Part

Our experimental scene setup is illuminated by a large spherical area light source surrounding the scene - the light probe. We get our light setup by sampling the light probe uniformly.

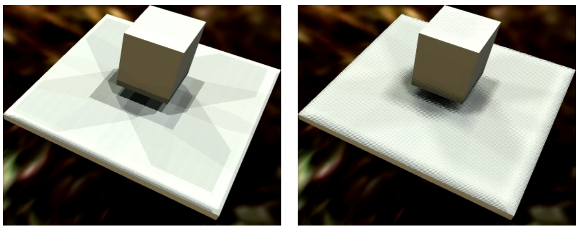

Figure 8: To smooth the transition between single shadows we introduce noise to the current sample position of the shadow map. Left image: shadow map sampling using only the map position, right image: shadow map sampling with noise, resulting in a smoother appearance.

These light sources are used in conjunction with shadow buffers to create the self shadowing effect. So we can control 2 parameters. The first is the number of light sources used; the second is the resolution of the shadow buffers used. To enhance the image quality of the borders between shadowed and non shadowed areas, we introduce noise to the shadow map sample position. (seen in figure 8.) The noise data is stored in a small tiled texture, which is repeated over the screen. This method introduces only low extra cost in difference to other smoothing or soft shadow methods. Additionally the shadow maps are applied in a screen space based method, which makes this part of the algorithm independent of the scene complexity. The cost of the shadowing process is only dependent of the number of shadows applied.

Local Part

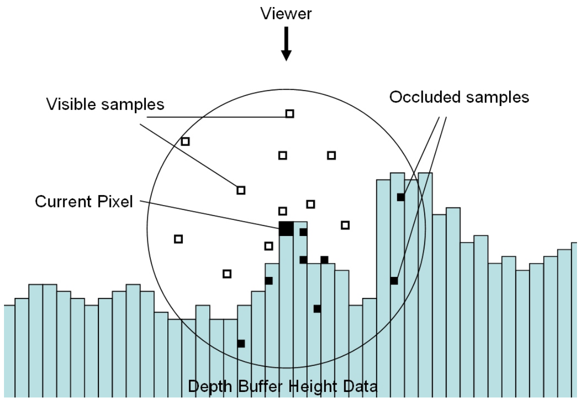

Figure 9: Self occlusion of the depth image, showing one line interpreted as height field: we test points inside a spherical region. The test results are related to the current sample point in the centre to get an approximation of its occlusion by its surroundings.

For the local self shadowing effects we have chosen a screen spaced ambient occlusion technique similar to the one presented in [SA07]. It only processes data of the depth component of the MRT buffer. In difference to the SSAO technique presented in [SA07] we used some simplifications. Our version samples the occlusion of the spherical area around the currently analysed pixel. We do not take into account the current normal. The distance attenuation of the importance of an occluder is automatically given by using a denser sampling towards the centre of the sphere. The amount of occlusion is simply approximated by averaging the test results of the sampled positions. Due to this sampling method, objects occluding less then the half of the sphere would appear brighter then usual. Since this is the case for all convex objects, we clamp the attenuation factor generated by the algorithm to the range between [0..1].

 ISBN 978-80-86943-94-7

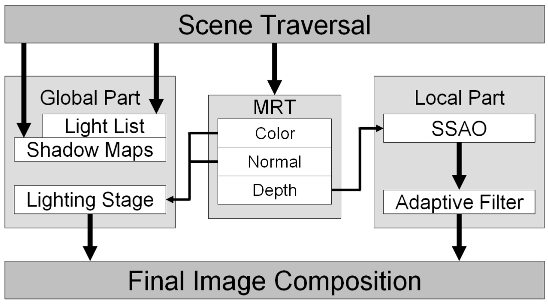

Figure 10: Overview of the used Rendering Architecture.

Architecture

The rendering system was implemented via a deferred shading approach. This allows a modular design and reduces the number of rendering passes for algorithms working on the current frame of the scene. An overview can be seen in figure 10. We start with a standard scene traversal and rendering of the scene. The result of the traversal is a light list, which is processed later in the pipeline. The scene is rendered to a multiple render target buffer wich holds the necessary information for the later deferred shading processes. In our case this buffer contains the colour component, the normal component and the depth component of the scene. We use the camera parameters to reconstruct the word coordinates of pixel data from the depth buffer information. The global pass is performed with the light list, an associated shadow map list and the colour + normal information from the MRT buffer. The local pass consists of the SSAO algorithm which works on the depth component of the MRT buffer. To reduce the number of texture reads inside the algorithm we make use of dithering. After the SSAO stage an adaptive filtering stage is applied to get a smooth appearance of the local pass output image. For the final combination the image from the lighting stage is attuned by the image from the local pass.

5. RESULTS AND DISCUSSION

We presented a hybrid technique for calculating the self shadowing of a fully dynamic scene with interactive to real-time frame rates. The technique splits the process into a local and a global method in order to combine the advantages of both. These methods can be chosen independently and are not limited to the ones we used. For the implementation of these two parts we presented a combination of a shadow buffer based self shadowing method and a screen spaces ambient occlusion algorithm. We have shown the capability of the local part to reduce light bleeding effects introduced by the method used in the global part. Additionally it adds shadowing details, which are not covered by the global part. Both parts are evaluated on the GPU using a deferred shading approach which is independent of the scene complexity. In our implementation we reach around 6 FPS when using an light probe approximation of 512 light sources and around 15 FPS when using an approximation of 128 lights.

6. FUTURE WORK

A benefit of the presented implementation is the possibility to use standard lighting models inside the global part. The presented screenshots use lambertian reflectance, but other reflection models can be used, too. A major problem of combining a directional occlusion and an ambient occlusion technique is over occlusion. Since both methods do not communicate with each other, light can be occluded several times. A future local approach should take into account the directional character of the incoming light. It would allow combining the directional occlusion information of both parts correctly. Additionally the lighting information of the light probe would also be taken into account in the local part. Another topic is the use of global illumination techniques inside the global and local part. The presented technique should be flexible enough to allow such modifications. However, an interesting aspect will be the combination of the local and the global part. With increasing complexity of the used algorithms it will be more and more necessary to exchange information between local and global part to get a seamless merge of the local and the global domain.

7. ACKNOWLEDGEMENTS

The high dynamic range environment maps used in this paper were made by Paul Debevec [Deb98].

8. REFERENCES

[CHL04] Greg Coombe, Mark J. Harris, and Anselmo Lastra. Radiosity on graphics hardware. In *GI '04: Proceedings of the 2004 conference on Graphics interface*, pages 161–168. Canadian Human-Computer Communications Society, 2004.

[Cro77] F.C. Crow. Shadow algorithms for computer graphics. *j-COMPGRAPHICS*, 11(2):242–248, July 1977.

[Deb98] Paul Debevec. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH '98: Proceedings of the 25th annual conference* *on Computer graphics and interactive techniques*, pages 189–198. ACM Press, 1998.

[DL06] William Donnelly and Andrew Lauritzen. Variance shadow maps. In *I3D '06: Proceedings of the 2006 symposium on Interactive 3D graphics and games*, pages 161–165, New York, NY, USA, 2006. ACM.

[DWS+88] Michael Deering, Stephanie Winner, Bic Schediwy, Chris Duffy, and Neil Hunt. The triangle processor and normal vector shader: a vlsi system for high performance graphics. In *SIGGRAPH '88: Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 21–30, New York, NY, USA, 1988. ACM.

[Gre86] Ned Greene. Environment mapping and other applications of world projections. *IEEE Comput. Graph. Appl.*, 6(11):21–29, 1986.

[HLHS03] Jean-Marc Hasenfratz, Marc Lapierre, Nicolas Holzschuch, and François Sillion. A survey of real-time soft shadows algorithms. In *Eurographics*. Eurographics, Eurographics, 2003. State-of-the-Art Report.

[HS99] Wolfgang Heidrich and Hans-Peter Seidel. Realistic, hardware-accelerated shading and lighting. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 171–178. ACM Press/Addison-Wesley Publishing Co., 1999.

[Kel97] Alexander Keller. Instant radiosity. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 49–56. ACM Press/Addison-Wesley Publishing Co., 1997.

[KF05] Martin Knuth and Arnulph Fuhrmann. Self-shadowing of dynamic scenes with environment maps using the gpu. In *WSCG'2005 Full Papers Conference Proceedings*, pages 31–38, 2005.

[KLA04] J. Kautz, J. Lehtinen, and T. Aila. Hemispherical rasterization for self-shadowing of dynamic objects. In *Proceedings Eurographics Symposium on Rendering 2004*, 2004.

[PBMH02] Timothy J. Purcell, Ian Buck, William R. Mark, and Pat Hanrahan. Ray tracing on programmable graphics hardware. *ACM Trans. Graph.*, 21(3):703–712, 2002.

[RH01] Ravi Ramamoorthi and Pat Hanrahan. An efficient representation for irradiance environment maps. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 497–500. ACM Press, 2001.

[RSC87] William T. Reeves, David H. Salesin, and Robert L. Cook. Rendering antialiased shadows with depth maps. *SIGGRAPH Comput. Graph.*, 21(4):283–291, 1987.

[SA07] Perumaal Shanmugam and Okan Arikan. Hardware accelerated ambient occlusion techniques on gpus. In *I3D '07: Proceedings of the 2007 symposium on Interactive 3D graphics and games*, pages 73–80, New York, NY, USA, 2007. ACM.

[SSZK04] M. Sattler, R. Sarlette, G. Zachmann, and R. Klein. Hardware-accelerated ambient occlusion computation. In B. Girod, M. Magnor, and H.-P. Seidel, editors, *Vision, Modeling, and Visualization 2004*, pages 331–338. Akademische Verlagsgesellschaft Aka GmbH, Berlin, November 2004.

[WBS03] Ingo Wald, Carsten Benthin, and Philipp Slusallek. Interactive global illumination in complex and highly occluded environments. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*, pages 74–81. Eurographics Association, 2003.

[Wil78] Lance Williams. Casting curved shadows on curved surfaces. In *SIGGRAPH '78: Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, pages 270–274. ACM Press, 1978.

[WPS+03] Ingo Wald, Timothy J. Purcell, Joerg Schmittler, Carsten Benthin, and Philipp Slusallek. Realtime ray tracing and its use for interactive global illumination. In *Eurographics State of the Art Reports*, 2003.

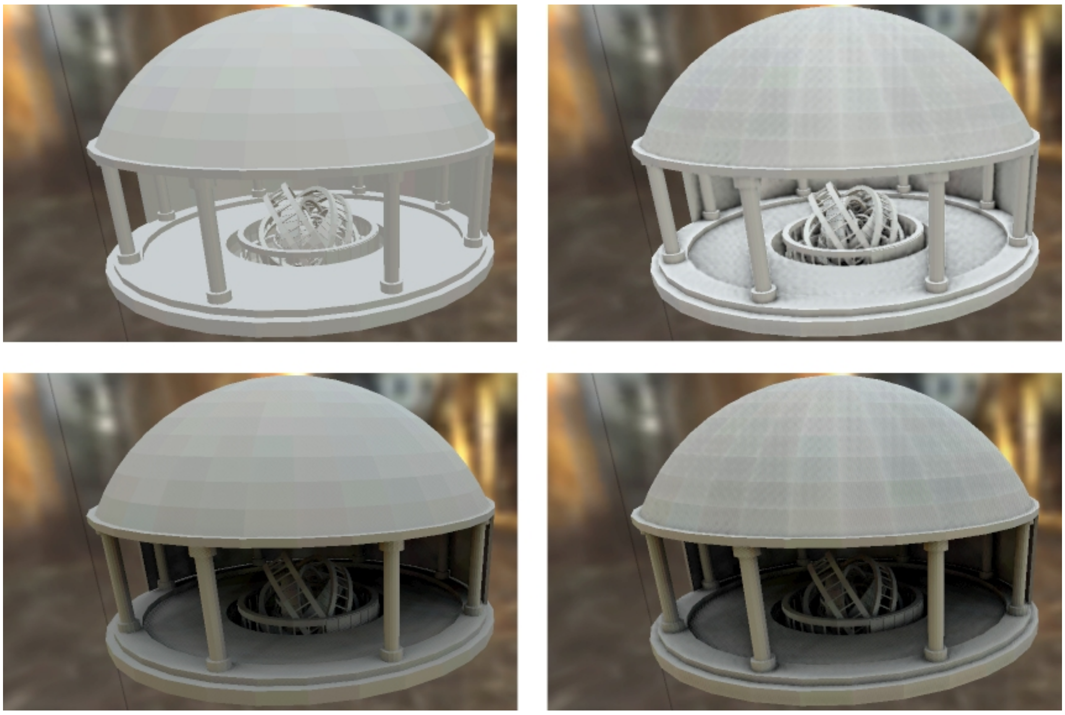

Figure 11: Comparison between the different effects of the presented technique. Upper left: no self shadowing. Upper right: local part. Lower left: global part. Lower right: combination of local and global part. The light probe was sampled using 512 light sources. The achieved an average of 6 fps.

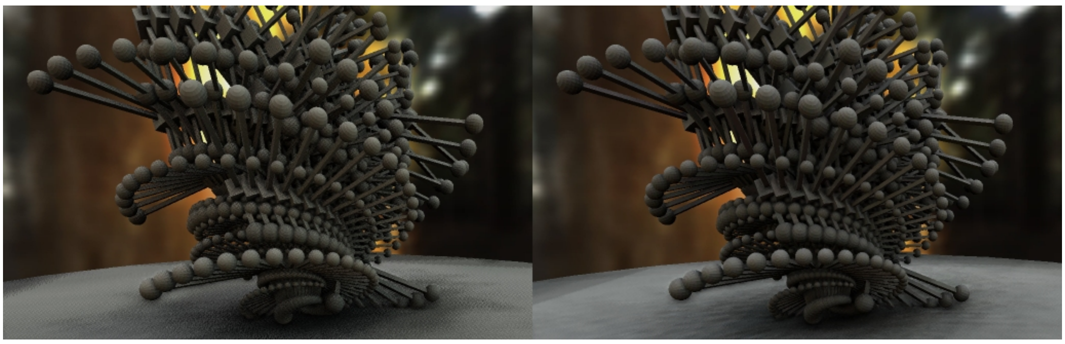

Figure 12: Comparison between different amounts of light sources in the global pass. Left image: 128 light sources at 15 fps. Right image: 512 light sources at 6 fps.

 ISBN 978-80-86943-94-7