**Horizon-Base Indirect Lighting**

## The ideal companion for your far-field indirect lighting solution.

February 2018 – Benoît “Patapom” Mayaux  
b.mayaux@gmail.com

In the following document, vectors will be written in bold letters (e.g. **,** ) and scalar values use standard letters (e.g. , ).  
The spherical coordinates used throughout the paper have an elevation angle aligned with the screen’s Z axis pointing toward the camera, and an azimuthal angle lying in the screen plane:

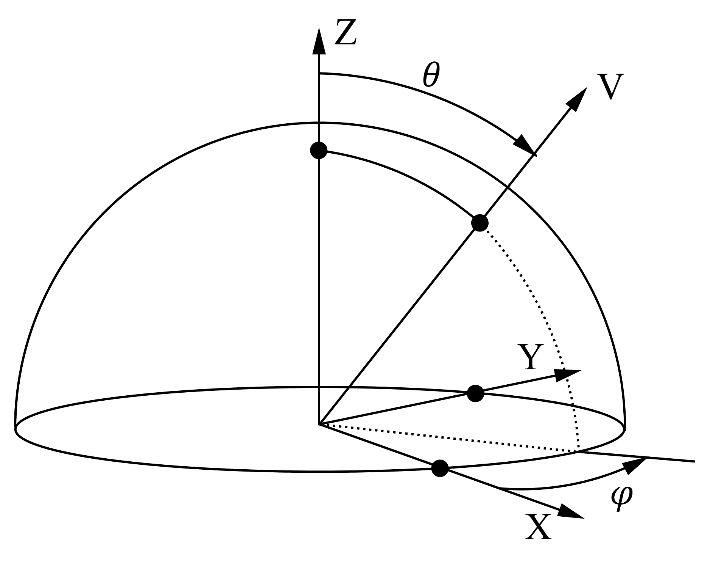


Fig. 1. The frame for spherical coordinates used in this paper.  
Z is pointing toward the camera. X and Y represent the image plane.

# Introduction

The Horizon-Based Ambient Occlusion (HBAO) technique was introduced by Bavoil et al. [[1](#REF_1)] and proposed to accelerate the computation of the Ambient Occlusion integral by skipping all the rays that we know would intersect the heighfield/depth buffer and thus wouldn’t contribute to the visibility term:

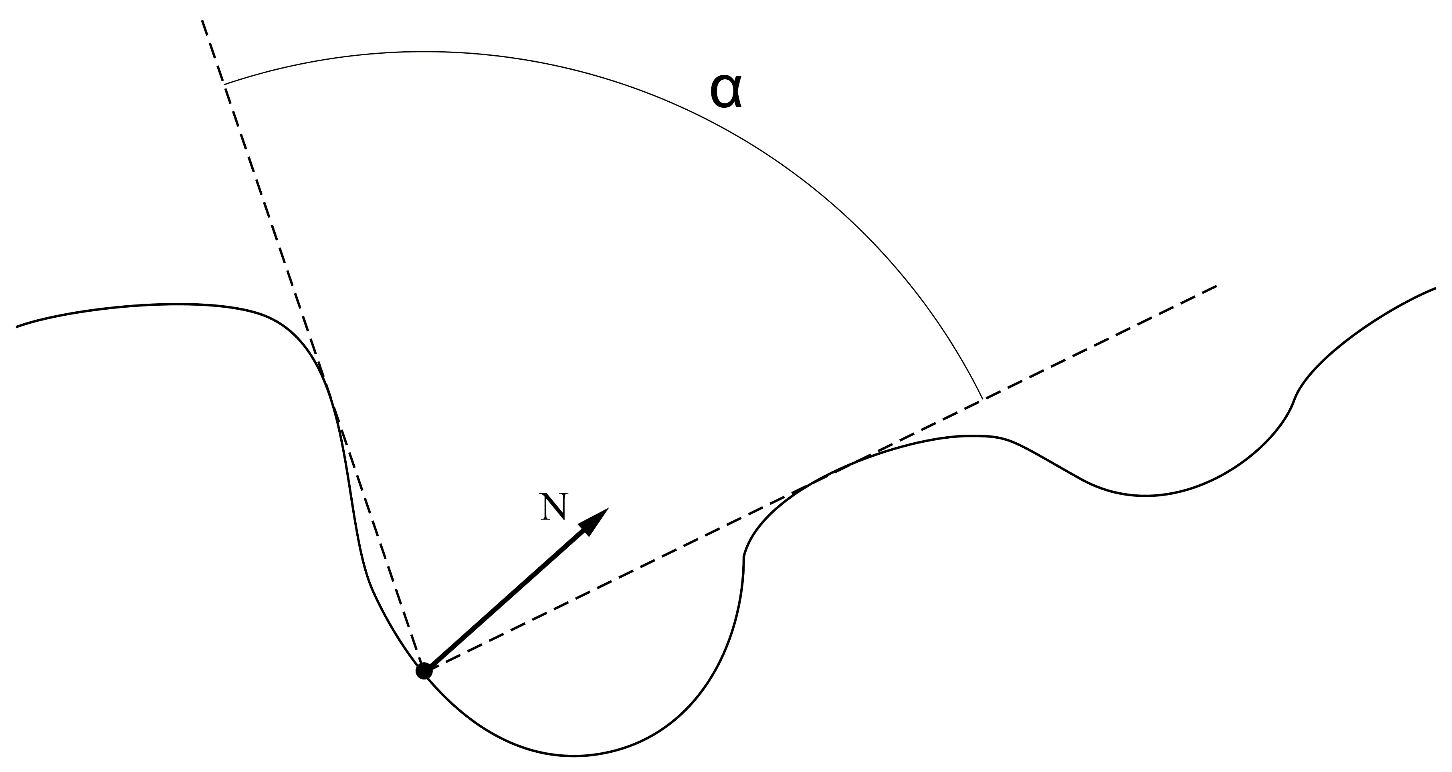


Fig. 2. The HBAO algorithm is exploiting the full visibility of the horizon cone of angle 𝛼:  
 we know all the rays in the cone escape the surface of the heightfield and are the only ones actually contributing to the visibility term.

The ambient occlusion term is then simplified into:

Where:

* is the location of current sample
* is the incoming ray direction
* is the upper hemisphere of directions whose solid angle is 2π
* is the portion of solid angle covered by the vector
* is the visibility term returning 1 if the ray escapes to infinity, 0 if the ray intersects the heightfield
* is the horizon angle in azimuthal direction

This technique allows us to avoid tracing rays for the entire hemisphere, instead we simply need to determine the horizon angle for a particular azimuthal angle by sampling the height field in screen space:

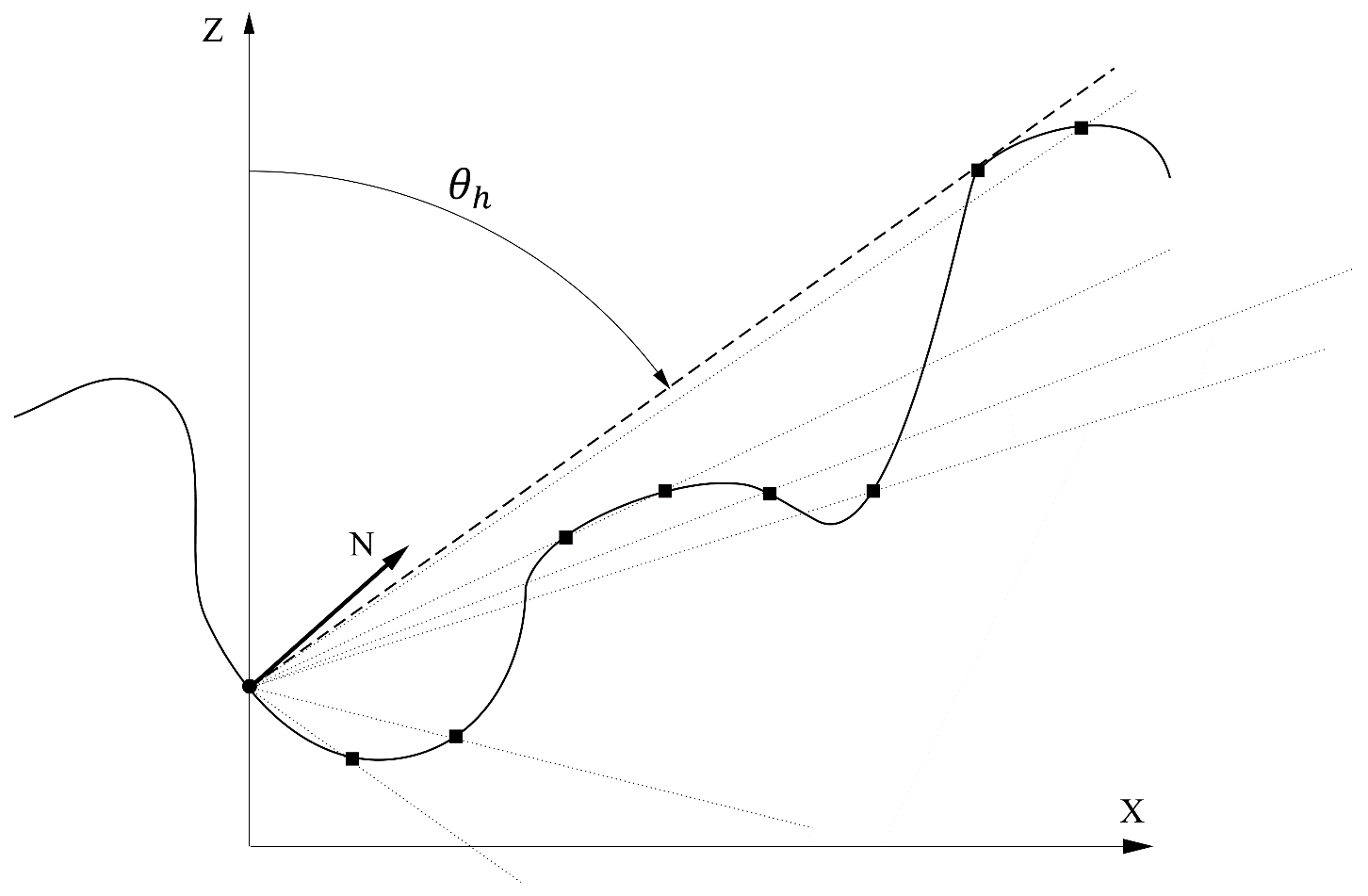


Fig. 3. The heightfield is sampled along the X direction and the horizon is updated along the way by reducing the horizon angle each time.

And finally, the portion of AO computed by the inner integral is:

And the final expression for the AO then becomes:

# Front & Back Sampling Convention

To avoid doing the same work twice, we will assume that we always trace the front and back slices of the screen-space disk at the same time:

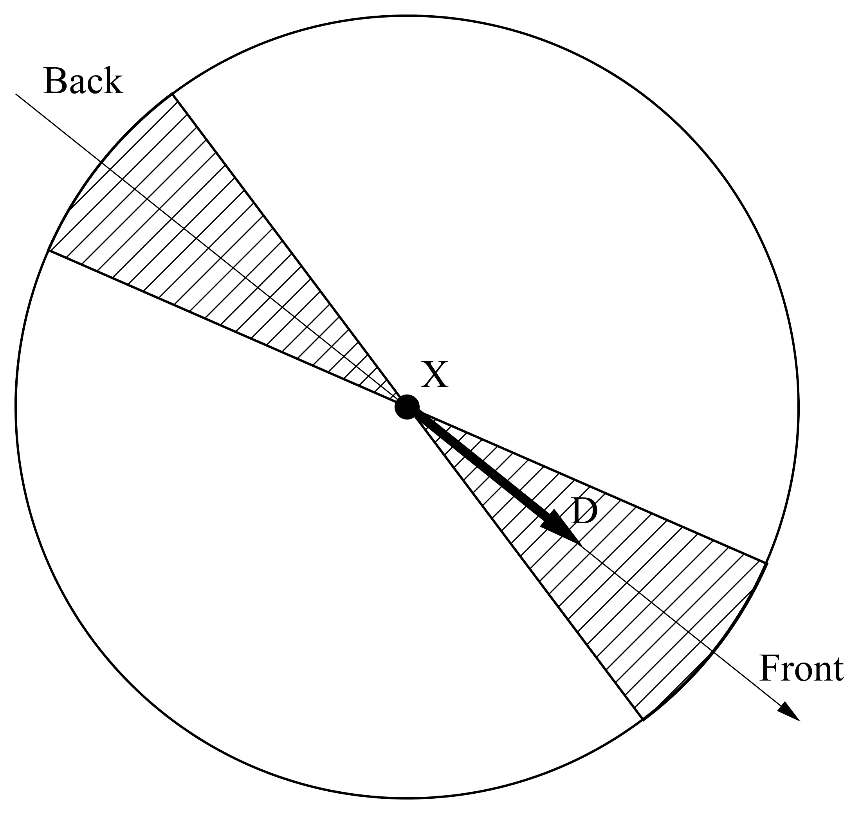


Fig. 4. The Screen-Space disk is sampled both in the forward and backward direction

Estimating the AO in such a simultaneous back & front configuration would then simply be:

# Using the Normal

We can first improve the quality of the HBAO algorithm by using the normal to initialize the horizon angles.  
The (camera-space) normal vector is often available from the G-Buffer produced by most deferred renderers.

In 3D camera-space, the gray areas of the 2D disk are approximated by a “slice” rotating about the Z axis:

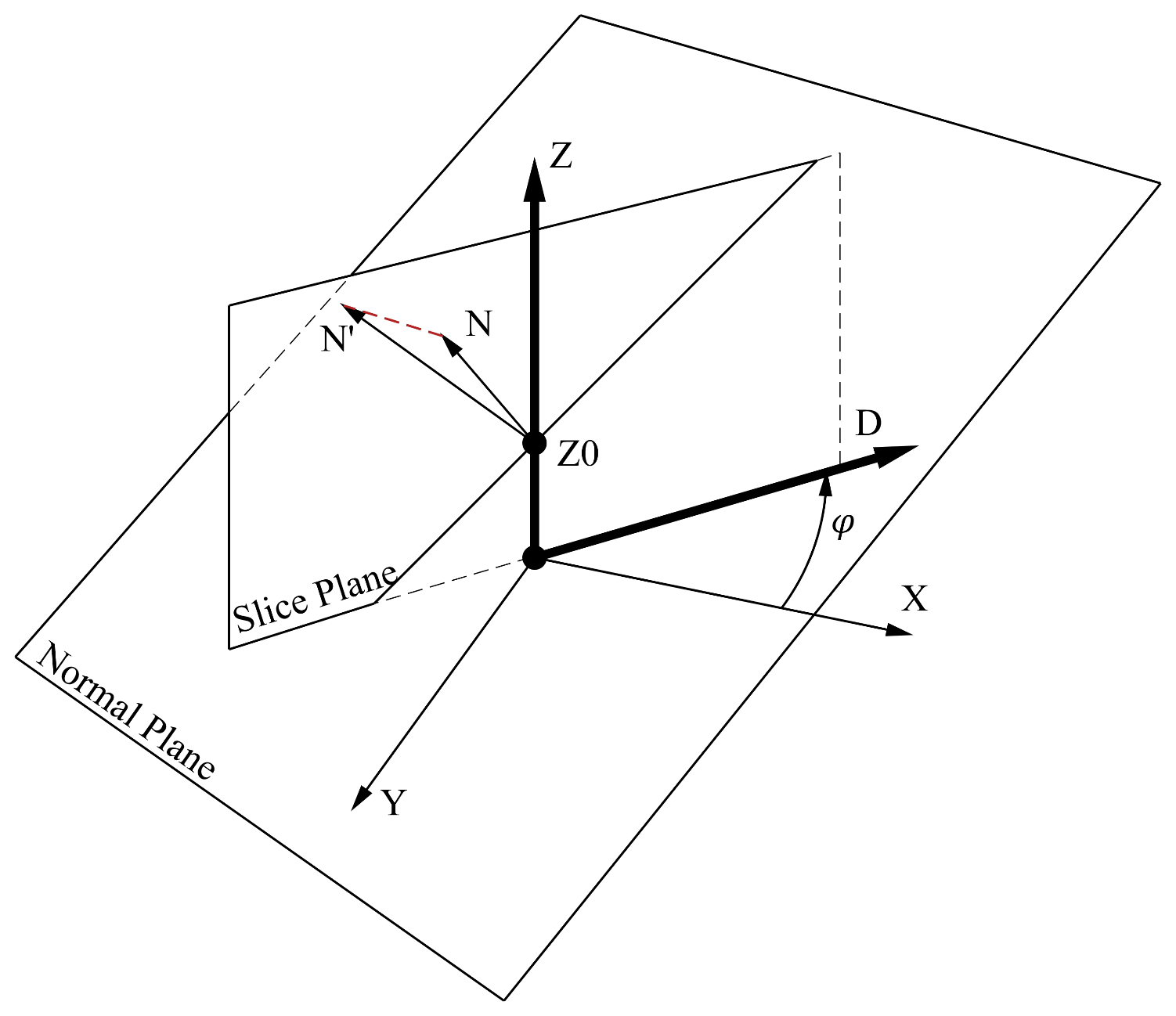


Fig. 5. a screen-space slice rotating by an angle about the Z axis pointing toward the camera. Z0 is the depth at our central location .  
We point out the important fact that the normal **N** is not necessarily lying in the plane of the slice and needs to be projected onto the slice, yielding the **N’** vector.

## Initial Horizon Angles

We initialize the horizon angles for the slice by computing the projection of the ***D*** vector onto the normal plane by following the **Z** axis:

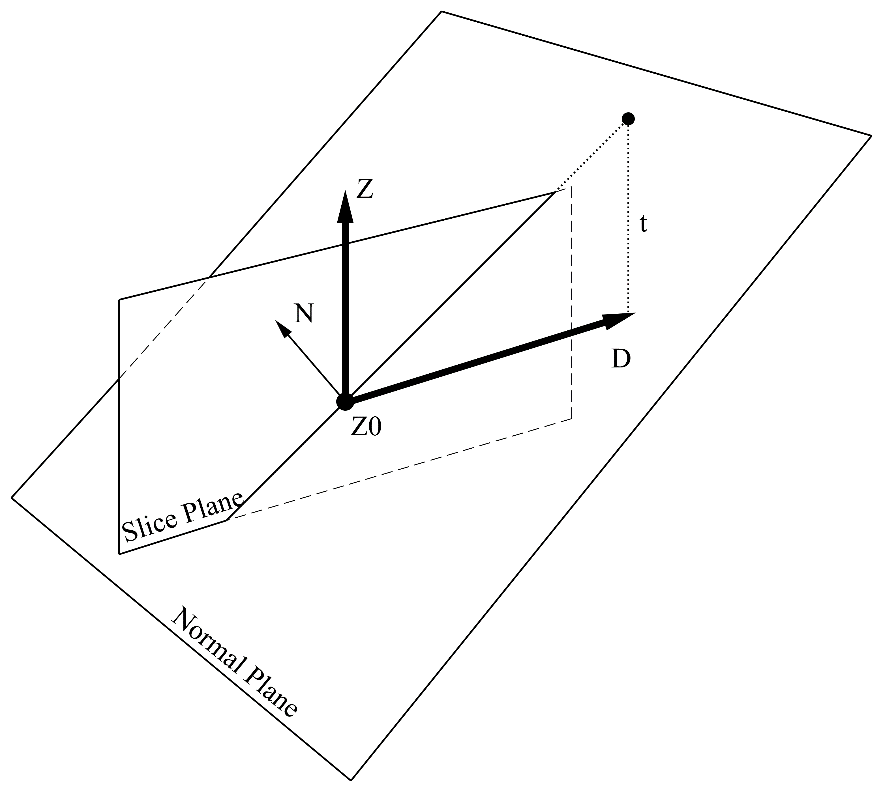


Fig. 6. Projection of the unit **D** vector onto the normal plane following the **Z** axis to find the initial “minimal” horizon angles for the slice.

Where:

* is the screen-space azimuthal direction 2D vector from figure 4.
* is the tangential part of the camera-space normal vector we sampled from the G-Buffer
* is the Z component of the camera-space normal vector we sampled from the G-Buffer
* is the intersection distance with the normal plane, in the 2D slice indicated by
* and are respectively the front and back horizon angles shown in figure 7.

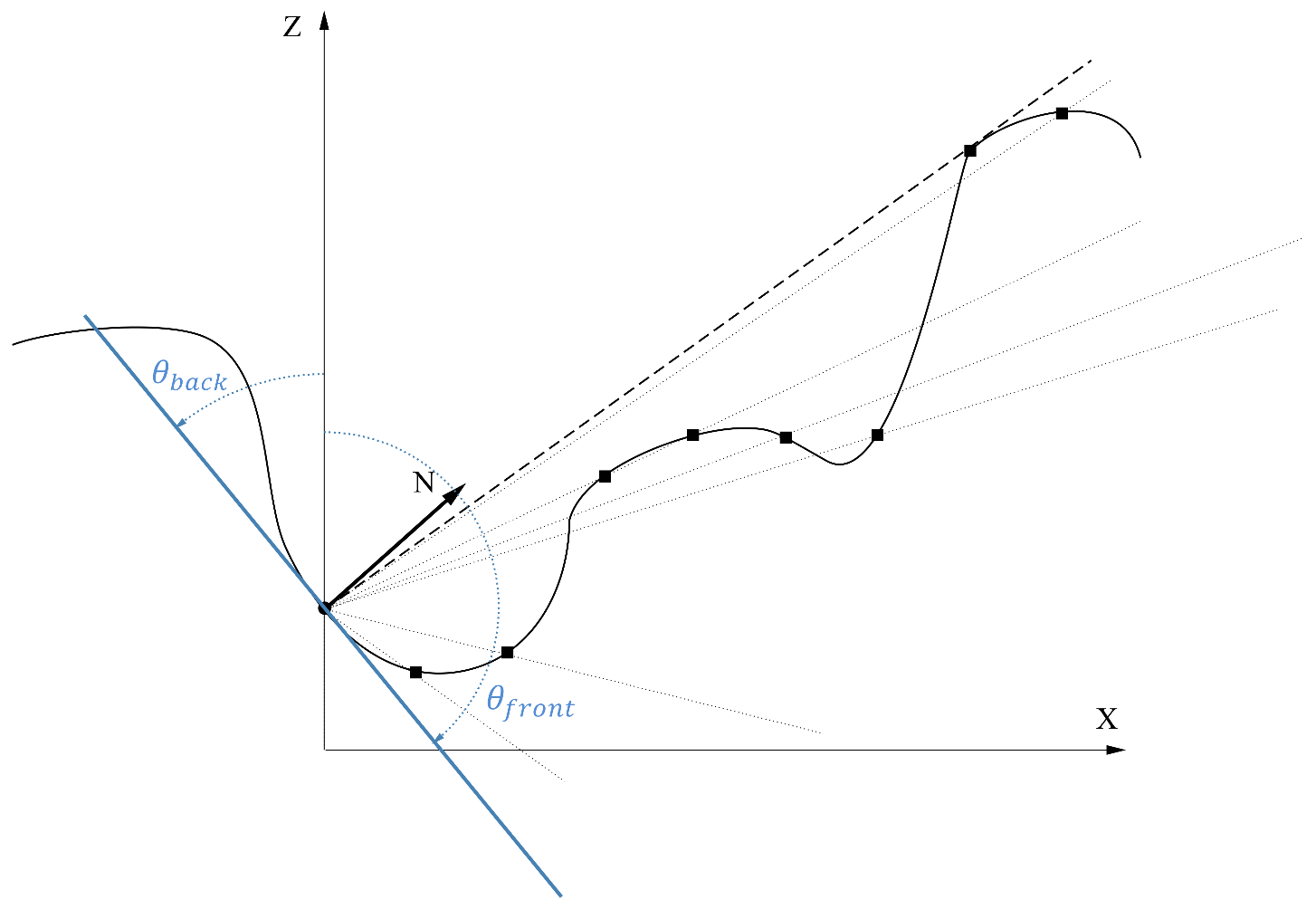


Fig. 7. Using the normal N to determine and as initial horizon values for each slice of the screen-space sampling disk.

## Updating the Horizon Angles

As we advance and sample along the front and back screen-space slice directions, we keep track of the maximum cosine of the angle with the Z axis as represented in figure 3 and 7:

Where:

* z is the z coordinate sampled from the height field / depth buffer
* is the z coordinate at the central location
* is the radius from the central location that increases as we take samples further and further away from
* is the measured elevation angle for the current sample
* is the horizon angle we keep updating as we move along

gna

# Bent Cones

Read & Credit Jimenez

# Indirect Lighting

The classical lighting equation to compute the outgoing radiance from a pixel at in viewing direction is essentially given by:

Where:

* is the incoming radiance at from direction
* is the surface’s BRDF
* is the surface normal
* is the set of all directions covering the upper hemisphere
* is the solid angle covered by the surface perceived along direction

Focusing only on diffuse Lambertian reflection, we can rewrite eq. (1) as:

Where:

* is the *irradiance* arriving at surface location and normal **.**
* represents the diffuse BRDF for a surface with albedo . The division by is here to guarantee energy conservation since .  
  NOTE: although a RGB quantity will be noted simply in the rest of the document

# Extracting the Common Ambient Occlusion Term

Introducing the visibility term:

We can rewrite eq. (2) as two distinct direct and indirect parts:

The direct term is (incorrectly) simplified into:

Where:

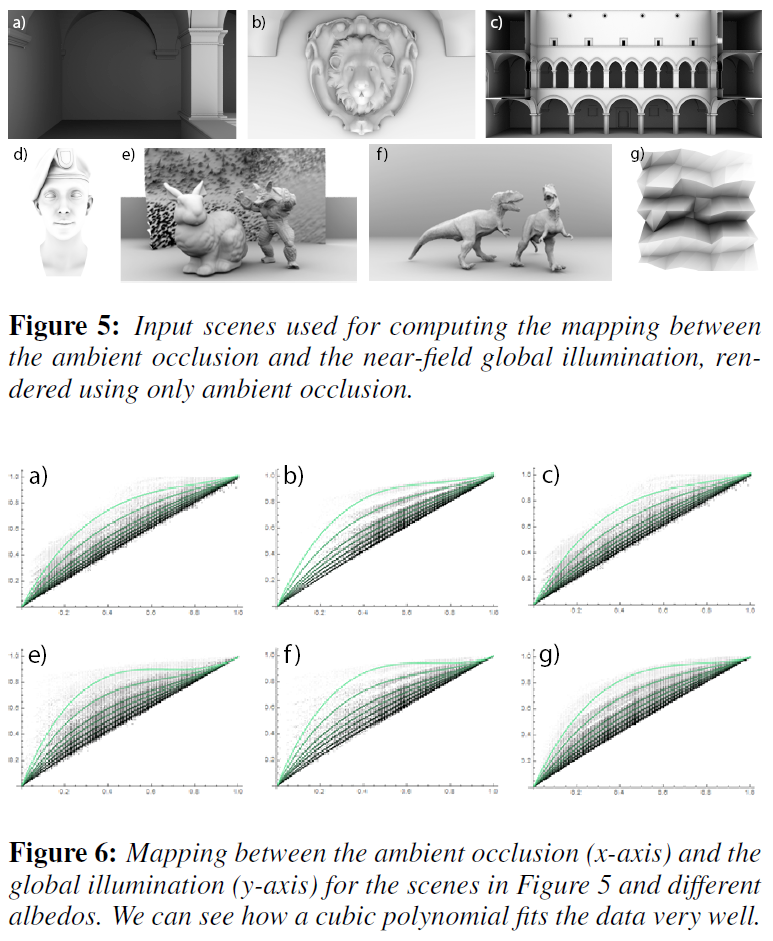
* is the directly-perceived (*i.e.* without any occlusion) irradiance estimate in the direction of the normal , which is usually pulled from a distant (in which case, the location is useless) encoding of the scene irradiance like a diffuse cube map, or spherical harmonics representation.
* is the scaled integration of the visibility term over all possible directions that we call the “Ambient Occlusion”. It’s a value we generally store per-vertex in meshes, or per-texel for AO texture maps.

We will see later how that simplification fares compared to the real deal.

# Indirect Term

The second indirect term is often ignored although it provides significant additional energy, especially when the visibility term is low (i.e. a very narrow aperture in the surface) and the surface reflectance is high (i.e. albedo close to 1).

Jimenez et al. gave a simplified value for the additional energy due to near-field inter-reflections in their 2016 paper *“*[*Practical Realtime Strategies for Accurate Indirect Occlusion*](http://iryoku.com/downloads/Practical-Realtime-Strategies-for-Accurate-Indirect-Occlusion.pdf)*”* (chapter 5) but they provided the energy “in bulk” by averaging empirical data from *only 3 bounces* of light, ignoring the subtle effects of each bounce on the albedo and the color saturation that ensues (unless their polynomial approximation is applied to each R, G, B component of the albedo independently, which is advised).



The curves showing the statistical relationship between AO value (x-axis) and the  
energy regained through near-field inter-reflections (y-axis).

First, let’s rewrite the indirect irradiance term replacing the reflected radiance using eq. (2):

Where:

* is the location and normal of the neighbor sample (*i.e.* indirect sample)
* is the irradiance perceived by the neighbor sample

To simplify the computation, we will pose that the surface reflectance at the neighbor site to be the same as the current location , which is a reasonable assumption for short distances, so we can write:

Since is the irradiance perceived by the neighbor surface in normal direction , we are introducing a recurrence relationship between the irradiance terms from successive bounces of light.

If we start from the initial directly-perceived irradiance term and assume ***unit radiance*** then:

Then the irradiance term perceived after a single bounce is:

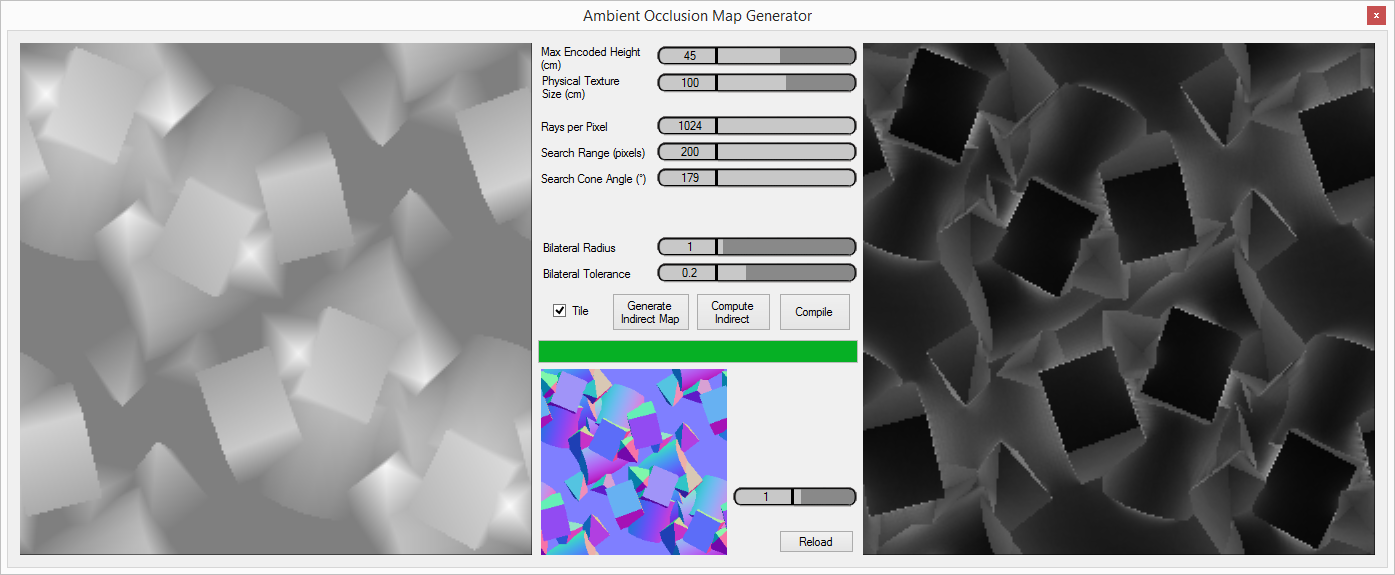
Writing the irradiance for yet another bounce gives:

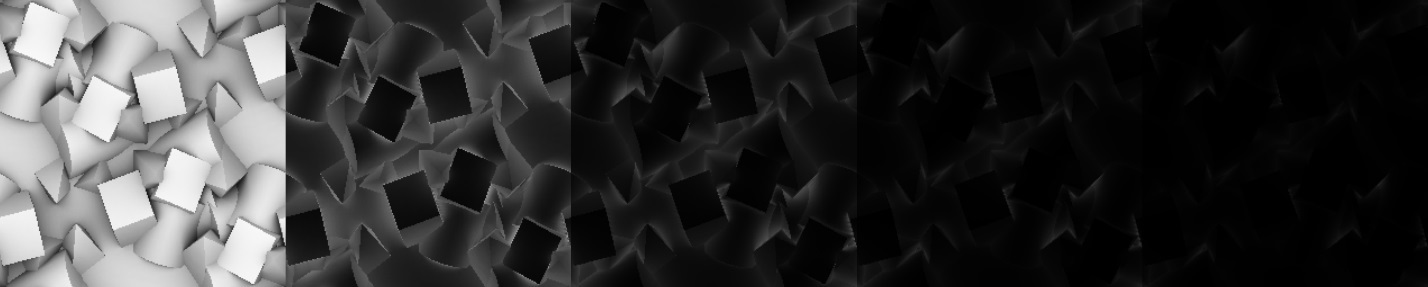
As a general rule, if we have a way to compute the integral then we can obtain:

We notice the subtle but interesting effect of color saturation induced by the raising to a power of the surface’s reflectance, which intuitively makes sense as a yellow wall will certainly get more “yellowy-orange” after a second bounce, even more after 3 bounces, etc. while absorbing energy every time a little more as well.

# Collecting Data

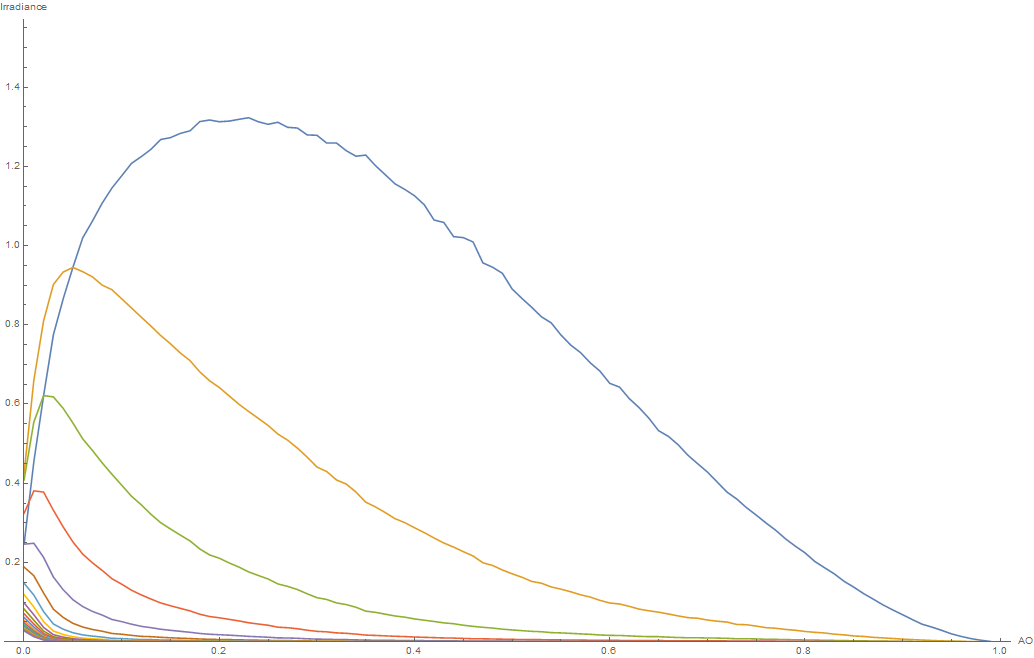
I didn’t want to write an entire path tracer like Jimenez did and instead decided to modify my little application that computes AO maps from height maps to also store “form factors” (a huge collection of links from each pixel to every visible neighbor) in order to perform the integration described above.





The effect of multiple bounces of light through a landscape.  
From left to right: directly-perceived irradiance, irradiance after 1 bounce, etc.

I collected irradiance data from various height map sources and finally ended up with curves for 20 distinct bounces:



The Irradiance Curves Obtained from the Simulation of 20 Bounces of Light

And represented as a continuous 3D graph:



Continuous representation of the decaying nature of light bounces

The algorithm to generate these curves goes like this:

1. For each pixel in the height map
   1. Ray-cast N rays to cover the hemisphere of direction and for each ray
      1. Accumulate 1 to the AO accumulator if the ray is unoccluded
      2. Accumulate to the direct radiance accumulator if the ray is unoccluded
      3. Add a new link to the perceived neighbor if the ray is occluded
   2. Normalize the accumulated data to obtain the AO and direct irradiance values
   3. Store into RG32F target
2. Initialize the source irradiance map with the direct irradiance that was just computed
3. For each bounce of light
   1. For each pixel in the height map
      1. For each occluded ray
         1. Accumulate neighbor’s irradiance \* (we use albedo=1)
      2. Normalize the accumulated data to obtain the reflected irradiance value
      3. Store into R32F target
4. Once AO texture, direct irradiance texture and all irradiance bounce textures are ready
   1. For each bounce
      1. For each pixel in the height map
         1. Accumulate the irradiance value into histogram bin whose index is selected based on the pixel’s AO
         2. Increment the bin’s counter by one
      2. Divide accumulated values in each bin by the counter
      3. Store histogram and display as curve

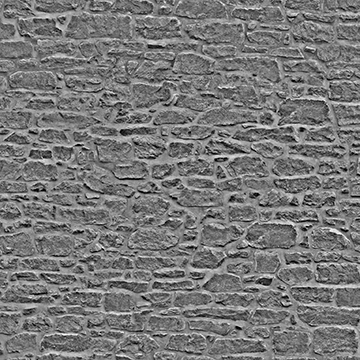
An implementation of the algorithm with compute shaders can be found at <https://github.com/Patapom/GodComplex/tree/master/Tests/TestGroundTruthAOFitting> .

## Effect of Maximum Height over the Curves

Varying the maximum amplitude of the various height maps didn’t have a lot of influence over the general aspect of the curves except for extreme values:

* Reaching extremely low amplitudes obviously prevented any bounce to occur and the curves had a tendency to flatten.
* Inversely, reaching very high amplitudes prevented light to reach the bottom of the surface, also flattening the curves.
* A good middle ground for these tests is to use height amplitudes that are roughly 1/10 the size of the map (so a 10cm height amplitude for a 1m large texture in my case).

Also, I found out it’s better to use quite “noisy” height map textures to obtain a lot of values for AO so each bin in the histogram has a significant number of samples.  
The best results were obtained using a highly detailed brick wall height map:



## Effect of the AO Integral Simplification over Actual Direct Illuminance

Armed with the direct and the multiple bounces irradiance curves, let’s first show the difference between actual luminance integration:

And the simplified AO model from eq. (4):

Where:

* , in that case, is the unoccluded irradiance from a diffuse cube map or some SH representation

Plotting the curve for a *unit* direct radiance , and the constant irradiance value of shows us an important term is missing:



Simplified AO Term (dashed blue) against Actual Irradiance (Red) for Unit Input Irradiance value

A good fit is found for the missing irradiance amplitude introducing a new mapping function :

With:

**

Matching the Actual Amplitude of the Irradiance

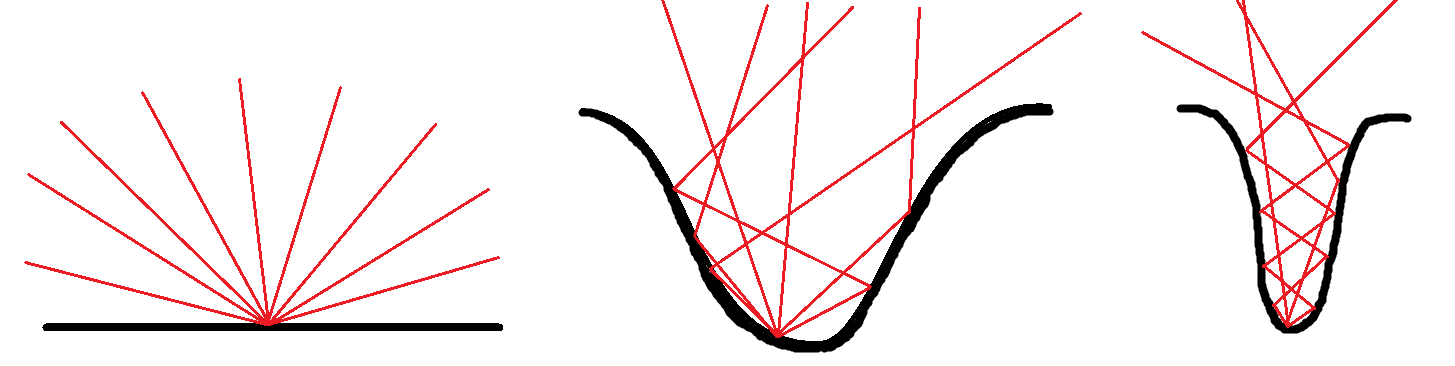
## Summing the Curves

Interestingly, summing the contribution of all the curves (including the direct irradiance curve) using a surface albedo and a constant yields a constant return of energy equal to π, whatever the AO value:



The Sum of All Bounces of Irradiance for a Constant Albedo of 100% = 𝜋

This result makes a lot of sense when you come to think of it: no energy can really be lost, the entirety of the energy is perceived either directly or indirectly through one or multiple bounces, regardless of the configuration of the geometric form factor:



For a 100% Reflecting Surface, the AO/Geometric Form Factor has  
 no Influence Over Total Perceived Irradiance

We could try and fit all the curves individually using some sort of log-normal distribution model:



Curve fitting for Individual Bounces

But at runtime, we are only given a single AO value and we want to find the irradiance values for each individual bounce.

We need to shift perspective by starting from the *initial irradiance value for a single-bounce* (i.e. the first, largest curve) for any given AO value, then we want to find the relationship between successive bounces so the irradiance value for the first bounce gives us the 2nd bounce, the 2nd bounce gives us the 3rd and so on.

Following the seemingly natural recurrence relationship between bounces from equation (5), let’s rewrite the expression for the final perceived irradiance after an infinite sum of bounces as:

Where:

* is the AO factor in [0,1]
* is the direct irradiance factor from equation (7)
* is the irradiance factor after 1 bounce
* is the irradiance factor after 2 bounces
* etc.

**NOTE**: I’m only using instead of here because the factor was already accounted for when the irradiance bounce curves were generated and fitted.

Moreover, we saw that when the initial irradiance and ρ=1:

And naturally:

Let's assume, for any AO value α and reflectance ρ=1, that we can rewrite , , (…) as a function of :

So:

And finally:

As seen on the 3D plot above, the amplitude of the bounces always decays more or less rapidly depending on so we can safely assume and then the geometric series can then be rewritten as:

Yielding:

Introducing the value of in the general expression of equation (8), we finally get:

We know the expression of from equation (7), all we need is to find , the factor to apply to to obtain the first bounce of irradiance…

## Finding

We saw earlier that we could obtain nice fits for the various bounce curves using a model like:

And indeed, a very good fit for is given by:

With:



Fitting for

# Runtime Usage

Let’s review and assemble everything!

We have the general indirect lighting expression that is given by eq. (2):

Then, from eq. (10) we have a way to decompose into our final irradiance approximation equation:

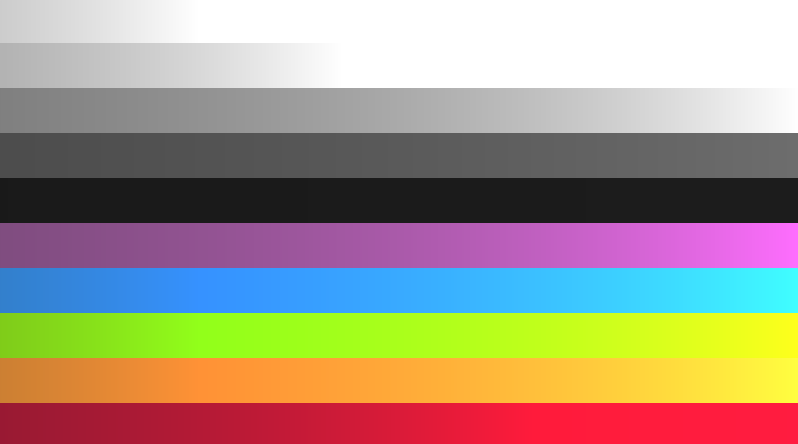
Using equations (7), (9) and (11) we have:

We can see below the final result of the fitting with a “unit” irradiance of for various albedo values:



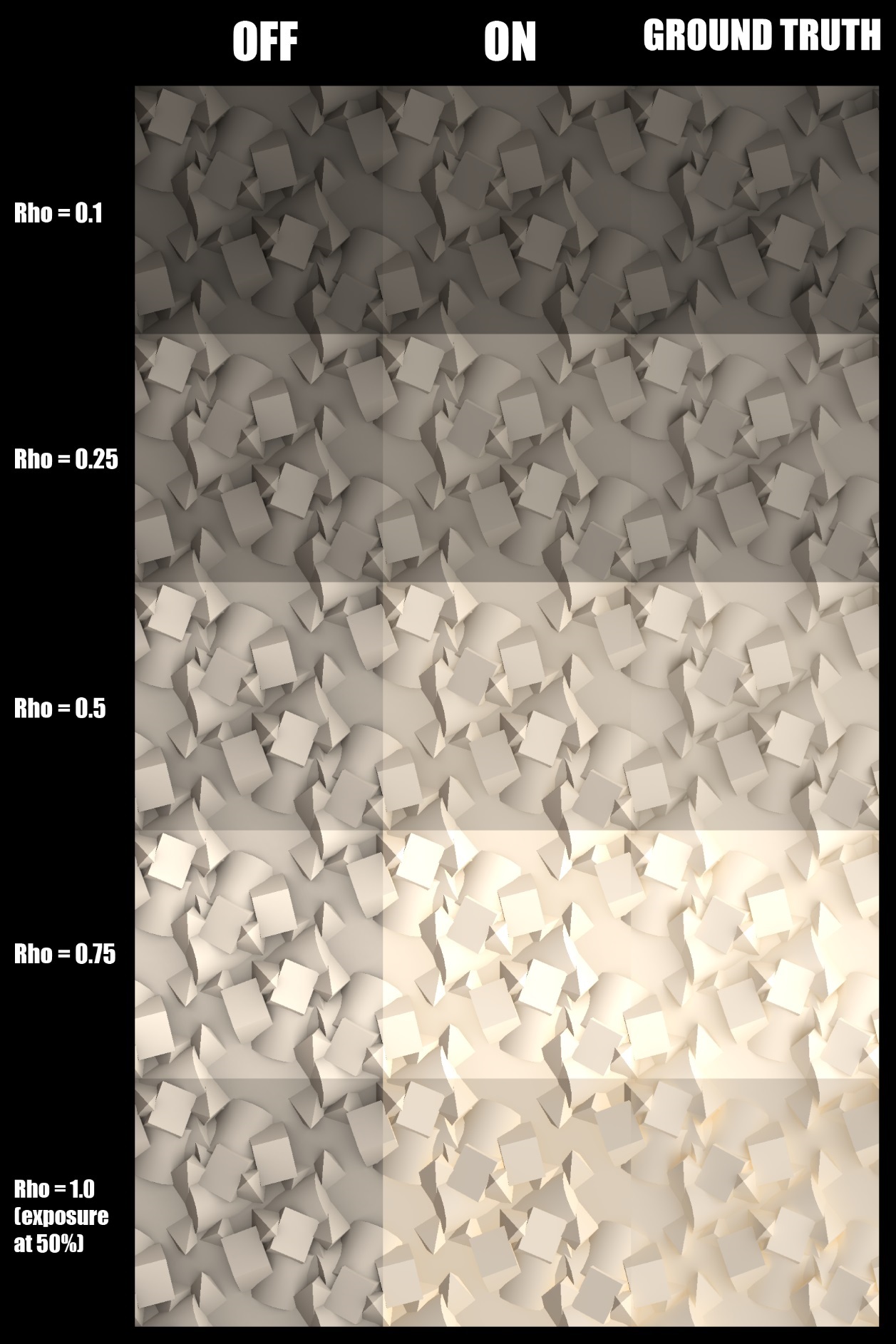
Resulting Curve Fitting for Various Values of from 0 to 100% every 10%.  
 Red: Experimental Data. Thick Curves: Fitted Model.

We should pay special attention to the part that needs to be executed as an (R,G,B) operation in order to show the nice color saturation that should naturally occur after multiple bounces of light:

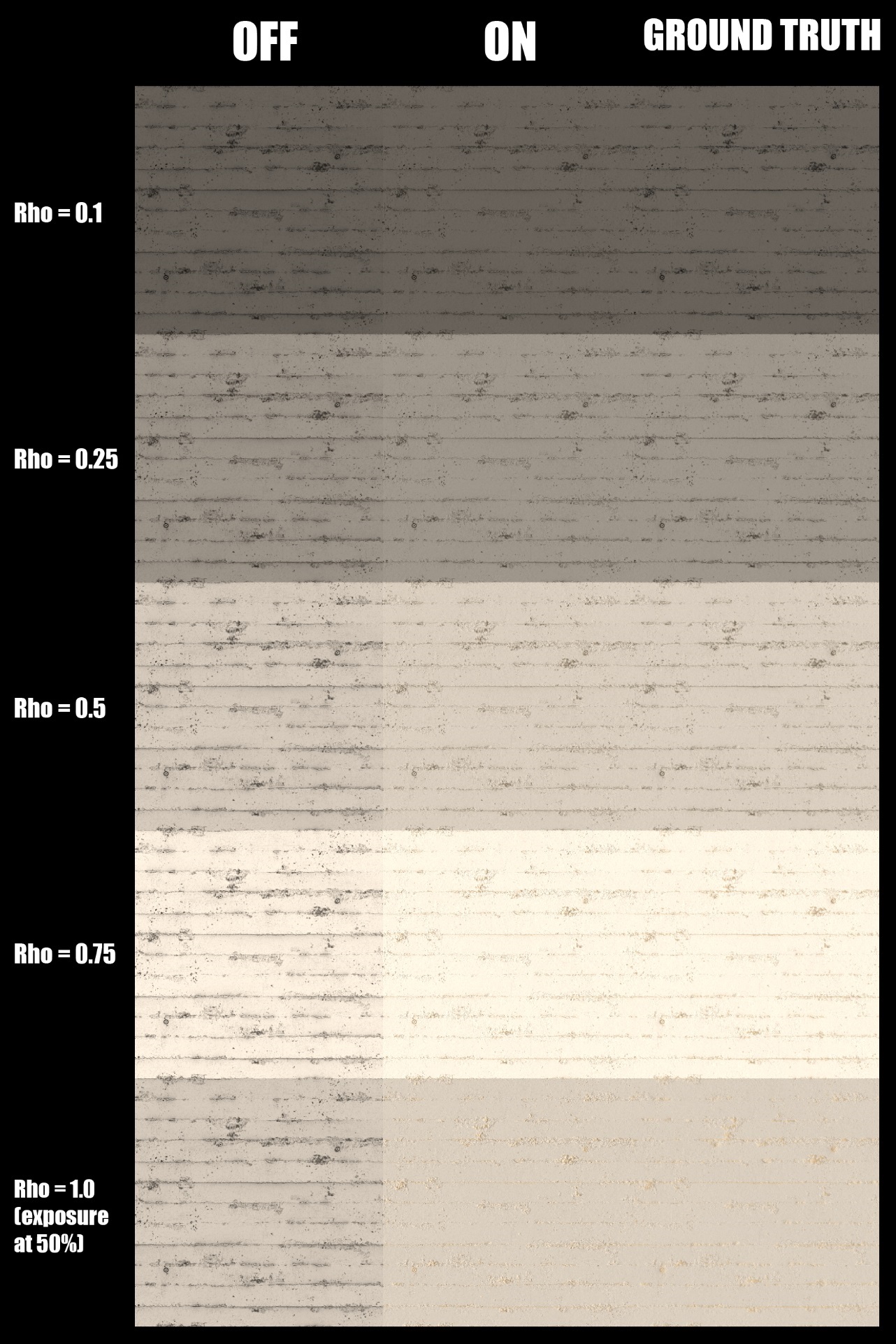


The Effect of Albedo Saturation for Varying Values of Tau from 0 to 1

# Results

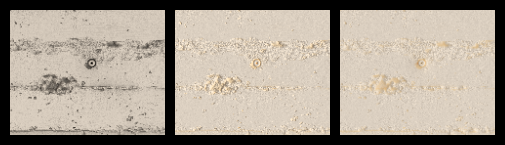


Geometric Shapes – 1m x 1m, 45cm Elevation



Concrete – 1m x 1m, 5cm Elevation





Details of the Subtle Color Bleeding Effect  
From Left to Right: Disabled, Enabled, Ground Truth

# Acknowledgments

Special thanks to Eric Arnebäck for proof reading this paper, Martin Gérard for his precious help with my math, [Geoffrey Rosin](https://www.artstation.com/kikette) for his amazing concrete texture, and Sandra for moral support 😊.

# References

*[1] Bavoil, L. and Sainz M. 2008. “Image-Space Horizon-Based Ambient Occlusion”*

*[2] Bavoil, L. and Jansen, J. 2013. “Particle Shadows & Cache-Efficient Post-Processing”*

*[3] Jimenez, J. Wu, X-C. Pesce, A. and Jarabo, A. 2016. “Practical Realtime Strategies for Accurate Indirect Occlusion”*

*[4] Cook, J. D. “*[*Accurately computing running variance*](https://www.johndcook.com/blog/standard_deviation/)*”*