

Study Notes:

Improved Ambient Occlusion

The classical lighting equation to compute the outgoing radiance from a pixel at \mathbf{x} in viewing direction $\boldsymbol{\omega}_o$ is essentially given by:

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) \rho(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \quad (1)$$

Where:

- $L_i(\mathbf{x}, \boldsymbol{\omega}_i)$ is the incoming radiance at \mathbf{x} from direction $\boldsymbol{\omega}_i$
- $\rho(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ is the surface's BRDF
- \mathbf{n} is the surface normal
- Ω^+ is the set of all directions covering the upper hemisphere
- $d\omega_i$ is the solid angle covered by the surface perceived along direction $\boldsymbol{\omega}_i$

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

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = \frac{\rho(\mathbf{x})}{\pi} \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i = \frac{\rho(\mathbf{x})}{\pi} E(\mathbf{x}, \mathbf{n}) \quad (2)$$

Where:

- $E(\mathbf{x}, \mathbf{n})$ is the *irradiance* arriving at surface location \mathbf{x} and normal \mathbf{n} .
- $\frac{\rho(\mathbf{x})}{\pi}$ represents the diffuse BRDF for a surface with albedo $\rho(\mathbf{x}) \in [0, 1]$. The division by π is here to guarantee energy conservation since $\int_{\Omega^+} (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i = \pi$.

Extracting the Common Ambient Occlusion Term

Introducing the visibility term:

$$V(\mathbf{x}, \boldsymbol{\omega}_i) = \begin{cases} 1 & \text{if the ray is unoccluded by the surface in direction } \boldsymbol{\omega}_i \\ 0 & \text{otherwise} \end{cases}$$

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

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = \frac{\rho(\mathbf{x})}{\pi} \left[\int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) V(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i + \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \right] \quad (3)$$

The direct term is (incorrectly) simplified into:

$$\int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) V(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\boldsymbol{\omega}_i \approx \left[\int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\boldsymbol{\omega}_i \right] \cdot \frac{\left[\int_{\Omega^+} V(\mathbf{x}, \boldsymbol{\omega}_i) d\boldsymbol{\omega}_i \right]}{2\pi} \approx E_0(\mathbf{x}, \mathbf{n}) \cdot AO(\mathbf{x}) \quad (4)$$

Where:

- $E_0(\mathbf{x}, \mathbf{n})$ is the directly perceived (*i.e.* without any occlusion) irradiance estimate in the direction of the normal \mathbf{n} , which is usually pulled from a distant (in which case, the location \mathbf{x} is useless) encoding of the scene irradiance like a diffuse cube map, or spherical harmonics representation.
- $AO(\mathbf{x})$ 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-textel 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*” (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).

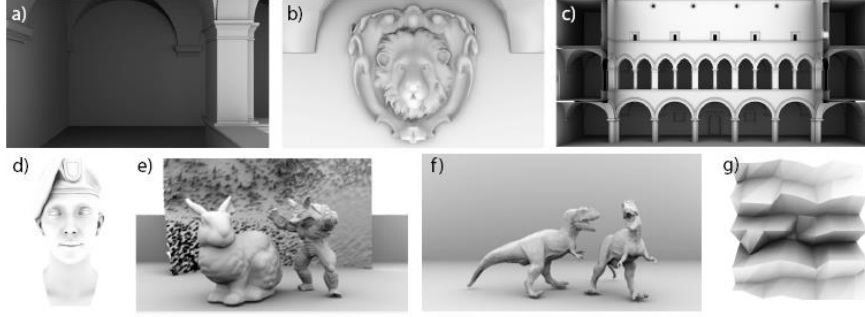


Figure 5: *Input scenes used for computing the mapping between the ambient occlusion and the near-field global illumination, rendered using only ambient occlusion.*

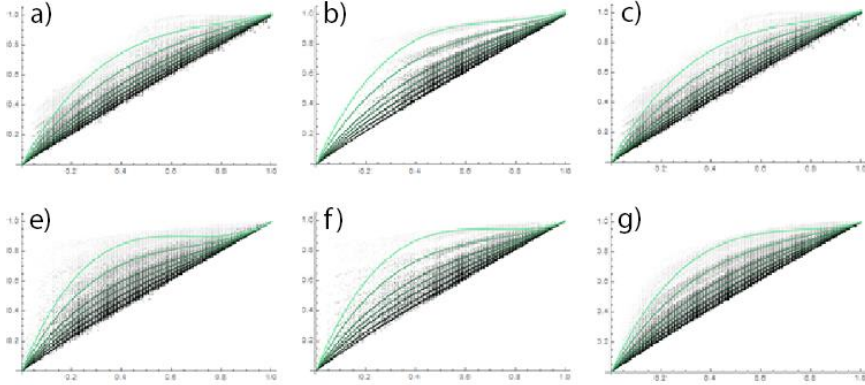


Figure 6: *Mapping between the ambient occlusion (x-axis) and the global illumination (y-axis) for the scenes in Figure 5 and different albedos. We can see how a cubic polynomial fits the data very well.*

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 $L_i(\mathbf{x}, \boldsymbol{\omega}_i)$ using eq. (2):

$$E(\mathbf{x}, \mathbf{n}) = \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\boldsymbol{\omega}_i = \int_{\Omega^+} \left[\frac{\rho(\mathbf{x}')}{\pi} E(\mathbf{x}', \mathbf{n}') \right] (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\boldsymbol{\omega}_i$$

To simplify the computation, we will pose that the surface reflectance at the neighbor site \mathbf{x}' to be the same as the current location $\frac{\rho(\mathbf{x}')}{\pi} = \frac{\rho(\mathbf{x})}{\pi}$ so we can write:

$$E(\mathbf{x}, \mathbf{n}) \approx \frac{\rho(\mathbf{x})}{\pi} \int_{\Omega^+} E(\mathbf{x}', \mathbf{n}') (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i$$

Where $E(\mathbf{x}', \mathbf{n}')$ is the irradiance perceived by the neighbor surface in normal direction \mathbf{n}' , introducing a recurrence relationship between the irradiance terms from successive bounces of light.

If we start from the initial irradiance term and assume **unit radiance** $L_i(\mathbf{x}, \boldsymbol{\omega}_i) = 1$ then:

$$E_0(\mathbf{x}, \mathbf{n}) = \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) V(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i = \int_{\Omega^+} V(\mathbf{x}, \boldsymbol{\omega}_i) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i$$

Then the irradiance term after a single bounce is:

$$\begin{aligned} E_1(\mathbf{x}, \mathbf{n}) &= \frac{\rho(\mathbf{x})}{\pi} \int_{\Omega^+} E_0(\mathbf{x}', \mathbf{n}') \cdot (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \\ &= \frac{\rho(\mathbf{x})}{\pi} I_0(\mathbf{x}, \mathbf{n}) \end{aligned}$$

Writing the irradiance for yet another bounce gives:

$$\begin{aligned} E_2(\mathbf{x}, \mathbf{n}) &= \frac{\rho(\mathbf{x})}{\pi} \int_{\Omega^+} E_1(\mathbf{x}', \mathbf{n}') \cdot (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \\ &= \frac{\rho(\mathbf{x})}{\pi} \int_{\Omega^+} \left[\frac{\rho(\mathbf{x})}{\pi} I_0(\mathbf{x}', \mathbf{n}') \right] \cdot (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \\ &= \left(\frac{\rho(\mathbf{x})}{\pi} \right)^2 \int_{\Omega^+} I_0(\mathbf{x}', \mathbf{n}') \cdot (1 - V(\mathbf{x}, \boldsymbol{\omega}_i)) (\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \\ &= \left(\frac{\rho(\mathbf{x})}{\pi} \right)^2 I_1(\mathbf{x}, \mathbf{n}) \end{aligned}$$

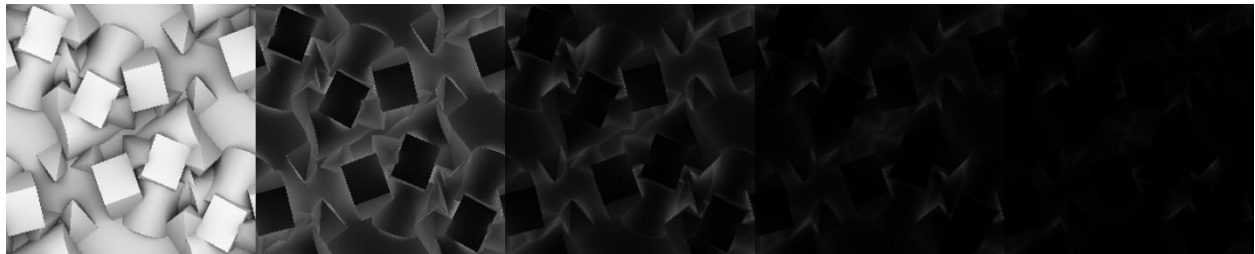
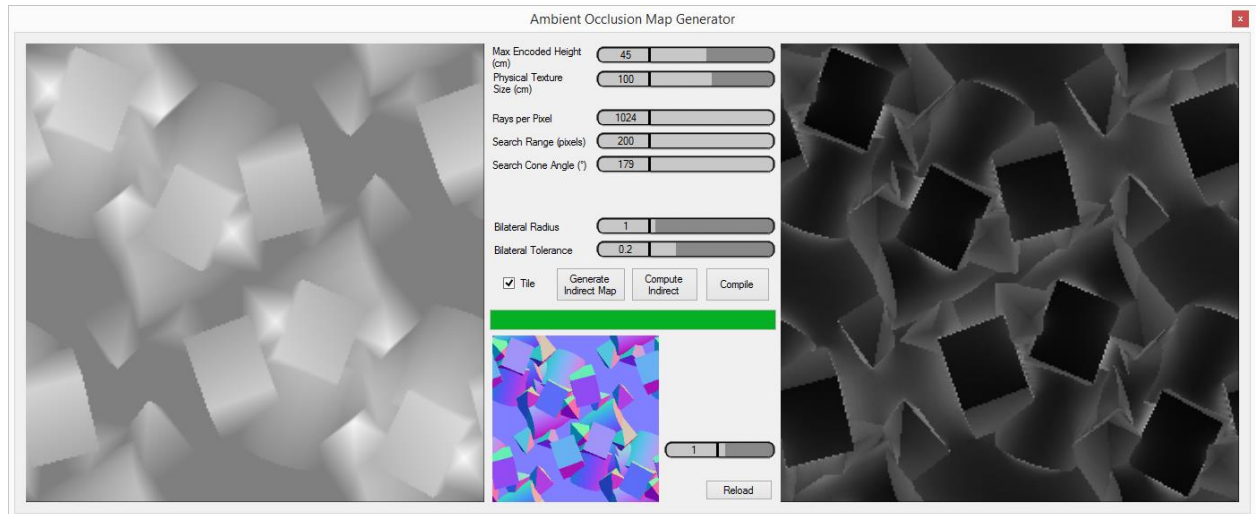
As a general rule, if have a way to compute the integral $I_{i-1}(\mathbf{x}, \mathbf{n})$ then we can obtain:

$$E_i(\mathbf{x}, \mathbf{n}) = \left(\frac{\rho(\mathbf{x})}{\pi} \right)^i I_{i-1}(\mathbf{x}, \mathbf{n})$$

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. and absorbing energy every time a little more.

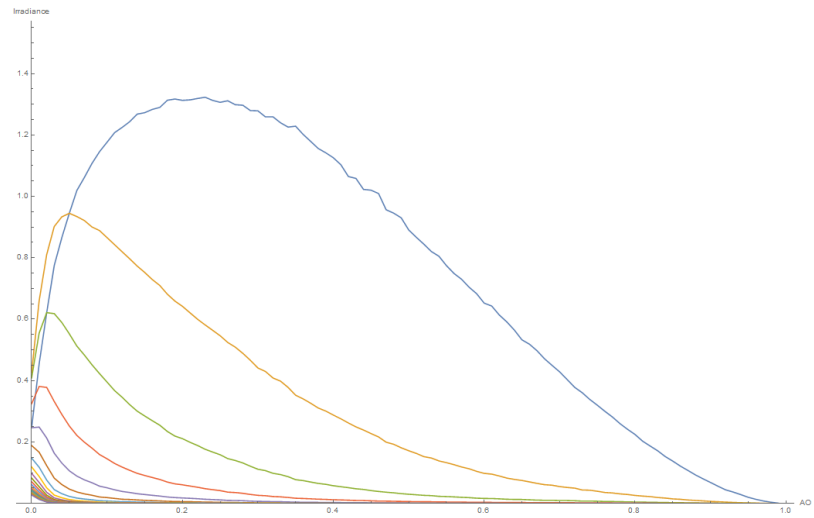
Collecting Data

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 each visible neighbor) in order to perform the integration described above.



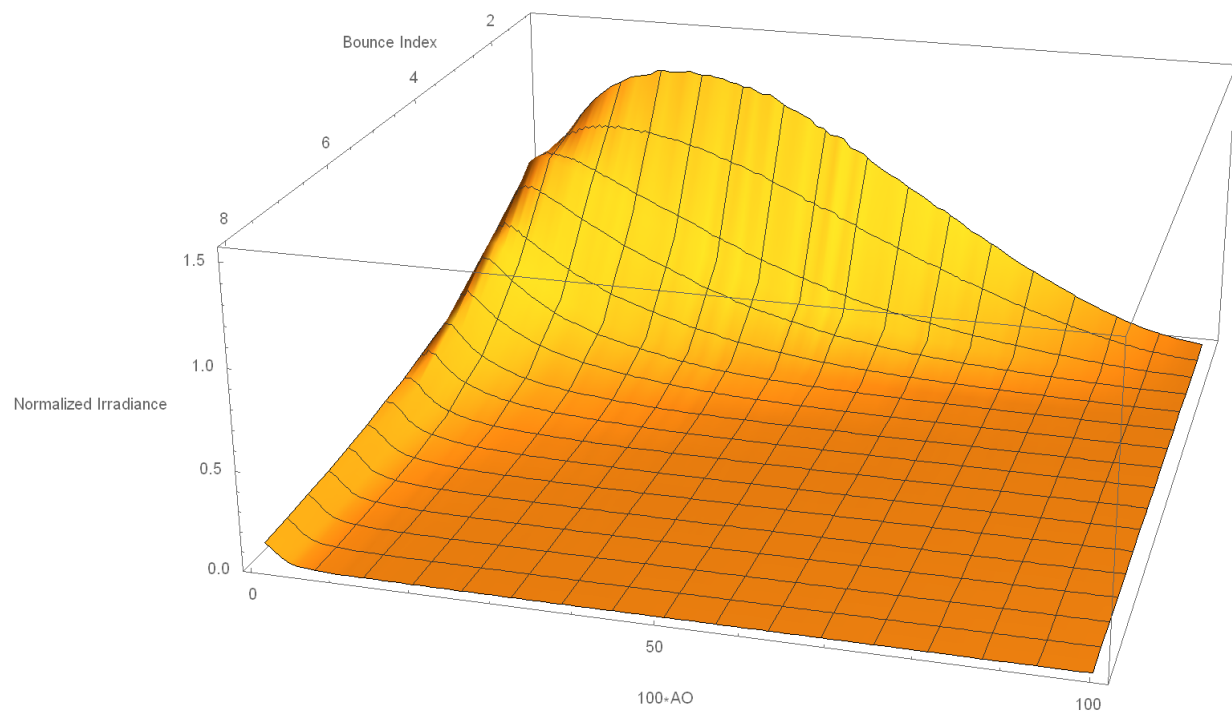
*THE EFFECT OF MULTIPLE BOUNCES OF LIGHT THROUGH A LANDSCAPE.
FROM LEFT TO RIGHT: DIRECTLY PERCEIVED IRRADIANCE, IRRADIANCE AFTER 1 BOUNCE, ETC.*

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
 - a. RAY-CAST N RAYS TO COVER THE HEMISPHERE OF DIRECTION AND FOR EACH RAY
 - i. ACCUMULATE 1 TO THE AO ACCUMULATOR IF THE RAY IS UNOCCLUDED
 - ii. ACCUMULATE $\cos(\theta)$ TO THE DIRECT RADIANCE ACCUMULATOR IF THE RAY IS UNOCCLUDED
 - iii. ADD A NEW LINK TO THE PERCEIVED NEIGHBOR IF THE RAY IS OCCLUDED
 - b. NORMALIZE THE ACCUMULATED DATA TO OBTAIN THE AO AND DIRECT IRRADIANCE VALUES
 - c. STORE INTO RG32F TARGET
- 2) INITIALIZE THE SOURCE IRRADIANCE MAP WITH THE DIRECT IRRADIANCE THAT WAS JUST COMPUTED
- 3) FOR EACH BOUNCE OF LIGHT
 - a. FOR EACH PIXEL IN THE HEIGHT MAP
 - i. FOR EACH OCCLUDED RAY
 1. ACCUMULATE NEIGHBOR'S IRRADIANCE * $\cos(\theta) * \frac{albedo}{\pi}$ (WE USE ALBEDO=1)
 - ii. NORMALIZE THE ACCUMULATED DATA TO OBTAIN THE REFLECTED IRRADIANCE VALUE
 - iii. STORE INTO R32F TARGET
- 4) ONCE AO TEXTURE, DIRECT IRRADIANCE TEXTURE AND ALL IRRADIANCE BOUNCE TEXTURES ARE READY
 - a. FOR EACH BOUNCE
 - i. 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
 - ii. DIVIDE ACCUMULATED VALUES IN EACH BIN BY THE COUNTER
 - iii. 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

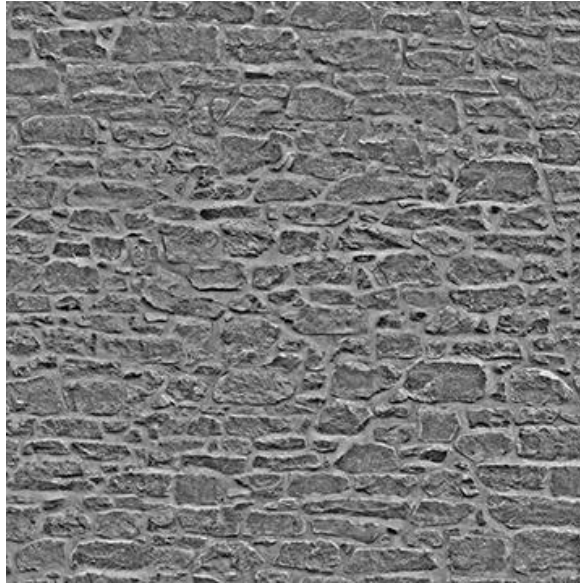
Varying the maximum amplitude of the various height maps didn't have a lot of influence over the general aspect of the curves except for extreme values:

- 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

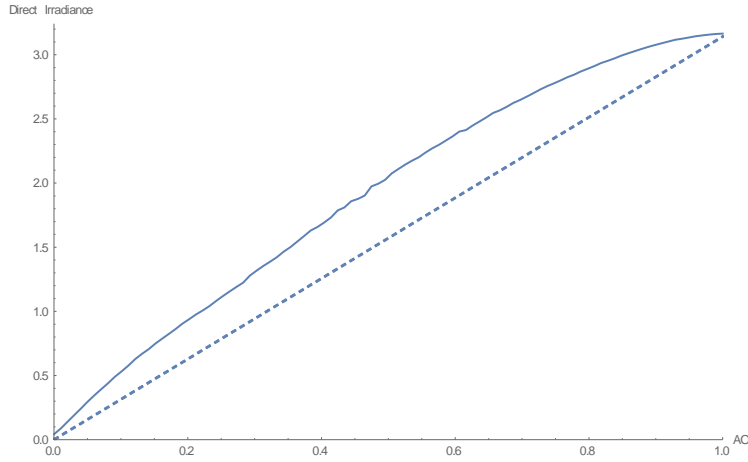
Armed with the direct and the multiple bounces irradiance curves, let's first show the difference between actual luminance integration:

$$E_0(\mathbf{x}, \mathbf{n}) = \int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i) \cdot V(\mathbf{x}, \boldsymbol{\omega}_i)(\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i$$

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

$$\tilde{E}_0(\mathbf{x}, \mathbf{n}) = \left[\int_{\Omega^+} L_i(\mathbf{x}, \boldsymbol{\omega}_i)(\mathbf{n} \cdot \boldsymbol{\omega}_i) d\omega_i \right] \cdot \left[\frac{\int_{\Omega^+} V(\mathbf{x}, \boldsymbol{\omega}_i) d\omega_i}{2\pi} \right] = E_0(\mathbf{x}, \mathbf{n}) \cdot AO(\mathbf{x})$$

Plotting the curve for a *unit* direct radiance $L_i(\mathbf{x}, \boldsymbol{\omega}_i) = 1$, and the constant irradiance value of $E_0(\mathbf{x}, \mathbf{n}) = \pi$ shows us an important term is missing:

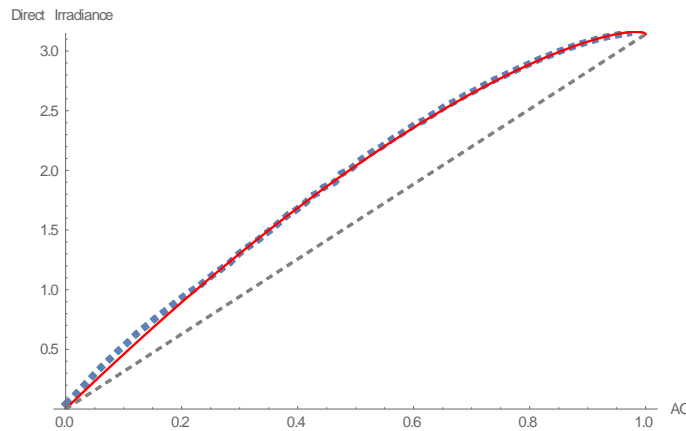


A good fit is found for the missing irradiance introducing a new mapping function $\mathcal{F}_0(\alpha)$:

$$\tilde{E}_0(\mathbf{x}, \mathbf{n}) = E_0(\mathbf{x}, \mathbf{n}) \cdot \mathcal{F}_0(AO(\mathbf{x})) \quad (5)$$

With:

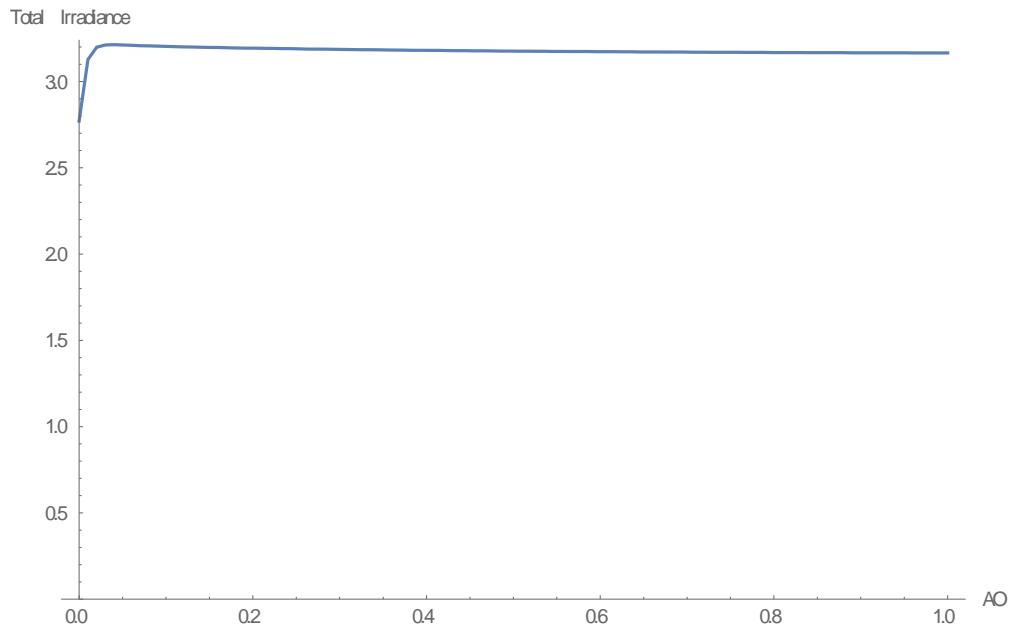
$$\mathcal{F}_0(\alpha) = \alpha * (1 + 0.5 * (1 - \alpha)^{0.75}) = \alpha * \left(1 + 0.5 * \frac{1 - \alpha}{\sqrt[4]{1 - \alpha}}\right) \quad (6)$$



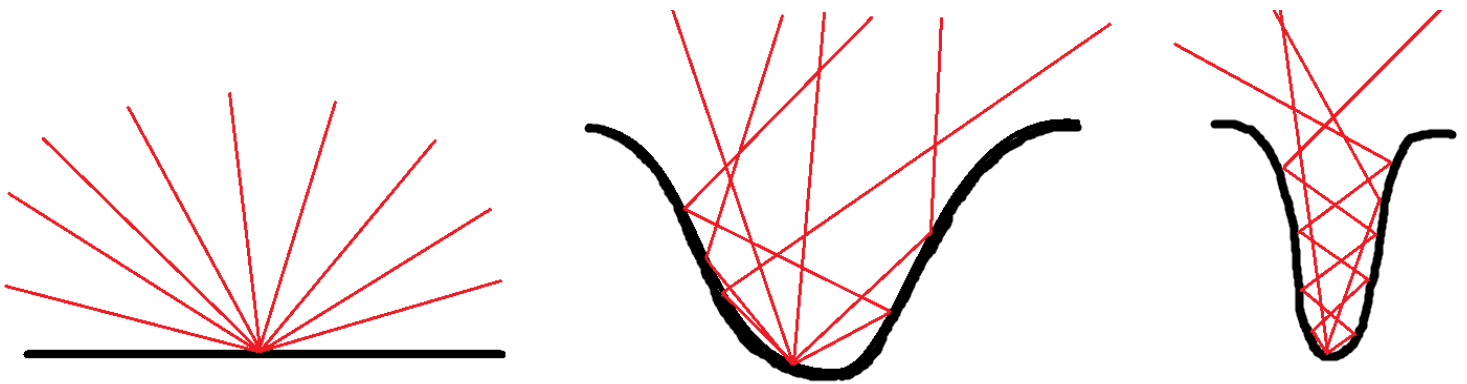
MATCHING THE ACTUAL AMPLITUDE OF THE IRRADIANCE

Summing the Curves

Interestingly, summing the contribution of all the curves (including the direct irradiance curve) using a surface albedo $\rho = 1$ and a constant $E_0(\mathbf{x}, \mathbf{n}) = \pi$ yields a constant return of energy equal to π , whatever the AO value:



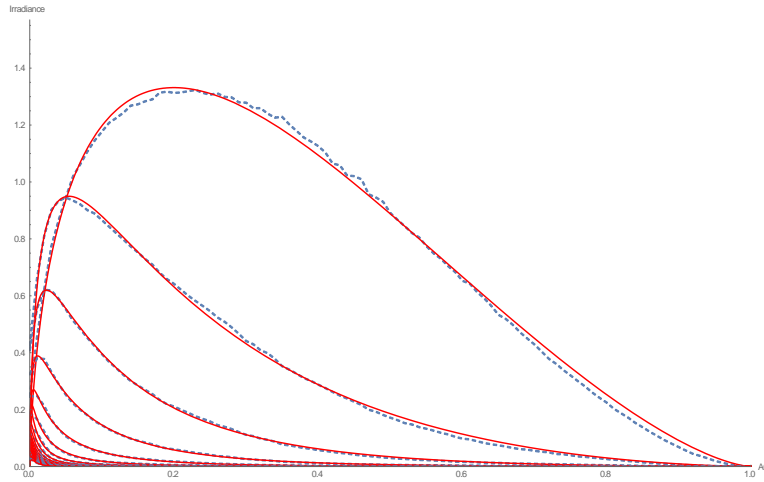
This result makes a lot of sense when you think of it: no energy can really be lost, the entirety of the energy is perceived either directly or indirectly through one or multiple bounces, regardless of the configuration of the geometric form factor:



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:

$$f(x) = a * x * (1 - x)^{1.5} * e^{-bx^c}$$



CURVE FITTING FOR INDIVIDUAL BOUNCES

But at runtime, we are only given a single AO value α 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.

Moreover, we are helped by the fact that the infinite sum of bounce values yields the initial irradiance $E_0(x, n) = \pi$ when $\rho=1$.

We can then assert that summing individual curves gives:

$$E_0(\alpha) + \rho E_1(\alpha) + \rho^2 E_2(\alpha) + \rho^3 E_3(\alpha) + \rho^4 E_4(\alpha) + (\dots) = \pi \quad (\text{for } \rho=1 \text{ and } E_0(x, n) = \pi)$$

Where:

α is the AO factor in $[0,1]$

$E_0(\alpha) = \pi \mathcal{F}_0(\alpha)$ is the irradiance directly perceived from the unoccluded environment

$E_i(\alpha) = \pi \mathcal{F}_i(\alpha)$ is the irradiance after i bounces

Let's assume, for any AO value α and reflectance $\rho=1$, that we can rewrite $E_2, E_3, (\dots)$ as a function of E_1 :

$$E_2(\alpha) = \tau E_1(\alpha)$$

$$E_3(\alpha) = \tau E_2(\alpha) = \tau^2 E_1(\alpha)$$

$$E_4(\alpha) = \tau E_3(\alpha) = \tau^3 E_1(\alpha)$$

$$(\dots)$$

So:

$$E_0(\alpha) + E_1(\alpha) + \tau E_1(\alpha) + \tau^2 E_1(\alpha) + \tau^3 E_1(\alpha) + (\dots) = \pi \quad (\text{for } \rho=1 \text{ and } E_0(\mathbf{x}, \mathbf{n}) = \pi)$$

$$E_0(\alpha) + E_1(\alpha) \sum_{b=0}^{\infty} \tau^b = \pi$$

And finally:

$$\sum_{b=0}^{\infty} \tau^b = \frac{\pi - E_0(\alpha)}{E_1(\alpha)}$$

As seen on the 3D plot above, the amplitude of the bounces always decays more or less rapidly depending on α so we can safely assume $\tau < 1$ and then the geometric series can then be rewritten as:

$$\frac{1}{1 - \tau} = \frac{\pi - E_0(\alpha)}{E_1(\alpha)}$$

Yielding:

$$\tau = 1 - \frac{E_1(\alpha)}{\pi - E_0(\alpha)} = 1 - \frac{\mathcal{F}_1(\alpha)}{1 - \mathcal{F}_0(\alpha)} \quad (7)$$

Applying the value of τ for $\rho = 1$ to the general expression, we finally get:

$$\begin{aligned} E(\alpha, \rho) &= E_0(\alpha) + \rho[E_1(\alpha) + \rho\tau E_1(\alpha) + (\rho\tau)^2 E_1(\alpha) + (\rho\tau)^3 E_1(\alpha) + \dots] \\ &= E_0(\alpha) + \frac{\rho}{1 - \rho\tau} E_1(\alpha) \quad (8) \end{aligned}$$

$E_0(\alpha)$ is given by eq. (5) taking $E_0(\mathbf{x}, \mathbf{n}) = \pi$.

All we need is to find $\mathcal{F}_1(\alpha)$, the factor to apply to $E_0(\mathbf{x}, \mathbf{n})$ to obtain the first bounce of irradiance...

Finding \mathcal{F}_1

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

$$f(x) = a * x * (1 - x)^{1.5} * e^{-bx^c}$$

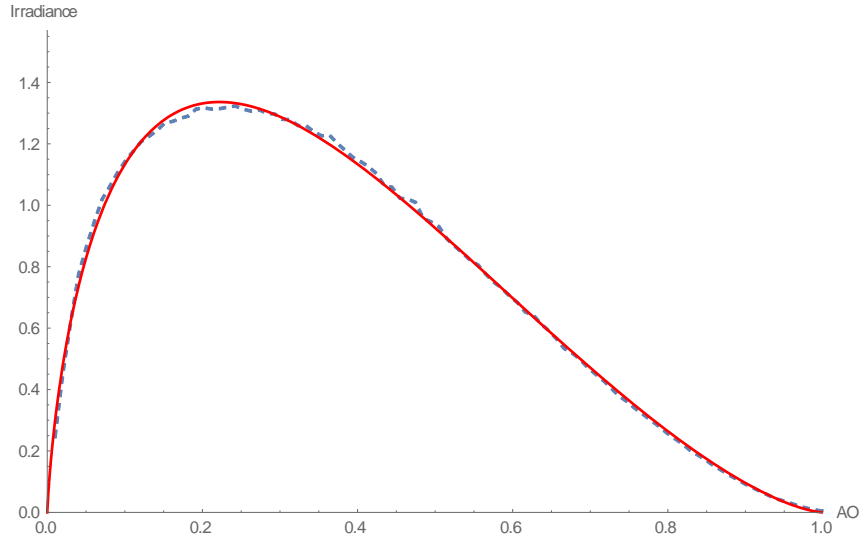
And indeed, a very good fit for $\mathcal{F}_1(\alpha)$ is given by:

$$\mathcal{F}_1(\alpha) = A. \alpha(1 - \alpha)^{1.5} e^{-B\sqrt[4]{\alpha}} \quad (9)$$

With:

$$A = 27.576937094210385$$

$$B = 3.3364392003423804$$



FITTING FOR $\mathcal{F}_1(\alpha)$

Runtime Usage

Let's review and assemble everything!

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

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = \frac{\rho(\mathbf{x})}{\pi} E(\mathbf{x}, \mathbf{n})$$

Then, from eq. (8) we have a way to decompose $E(\mathbf{x}, \mathbf{n})$ into our final irradiance approximation equation:

$$E(\mathbf{x}, \mathbf{n}) \approx E_0(\mathbf{x}, \mathbf{n}) \left(\mathcal{F}_0(AO(\mathbf{x})) + \frac{\rho(\mathbf{x})}{1 - \rho(\mathbf{x}) \cdot \tau} \mathcal{F}_1(AO(\mathbf{x})) \right) \quad (10)$$

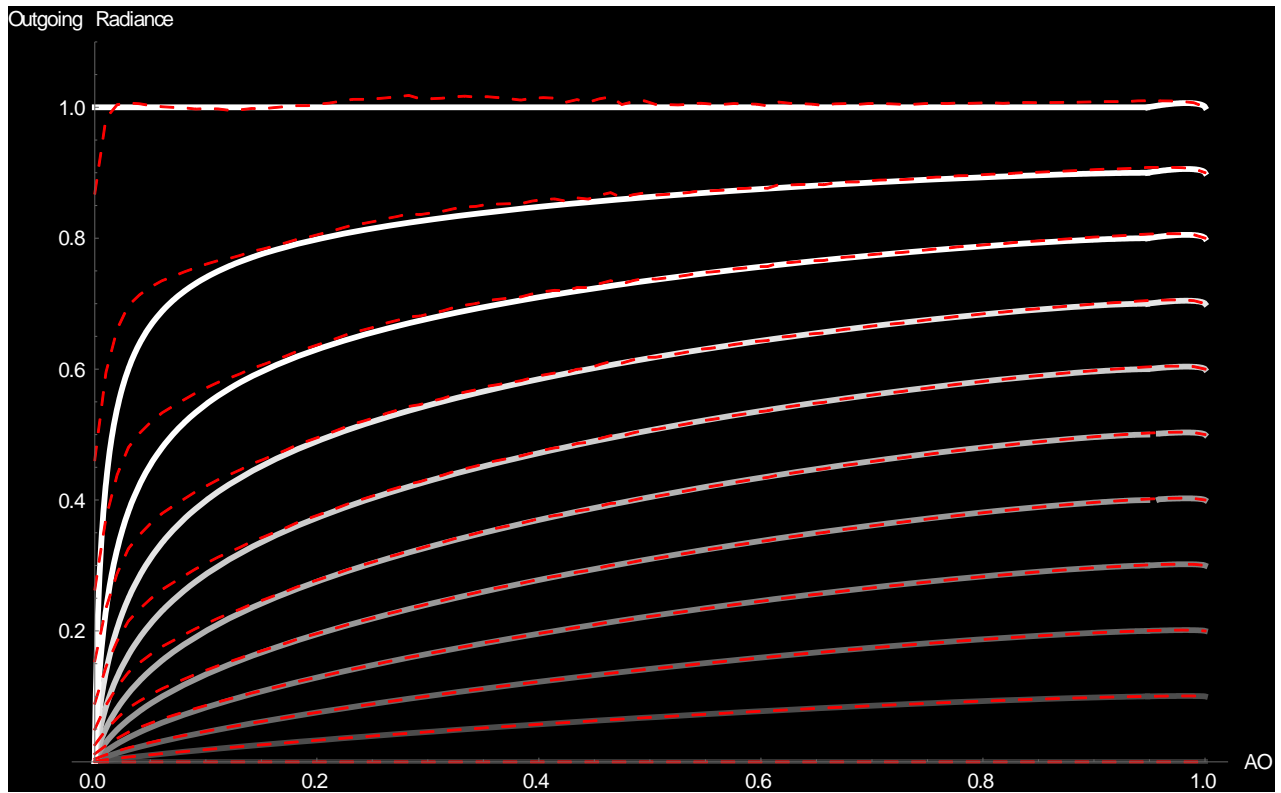
Using equations (6), (7) and (9) we have:

$$\mathcal{F}_0(\alpha) = \alpha \left(1 + \frac{(1 - \alpha)^{0.75}}{2} \right)$$

$$\mathcal{F}_1(\alpha) = A \cdot \alpha (1 - \alpha)^{1.5} e^{-B \sqrt[4]{\alpha}}$$

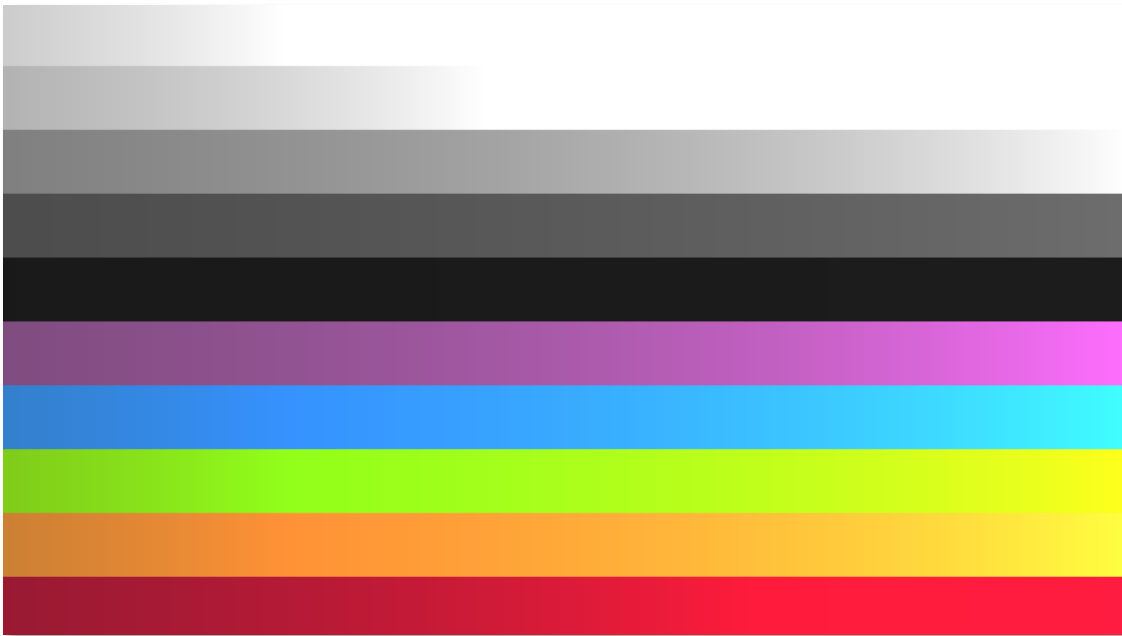
$$\tau = 1 - \frac{\mathcal{F}_1(\alpha)}{1 - \mathcal{F}_0(\alpha)}$$

We can see below the final result of the fitting with a “unit” irradiance of $E_0(\mathbf{x}, \mathbf{n}) = \pi$ for various values of albedo:



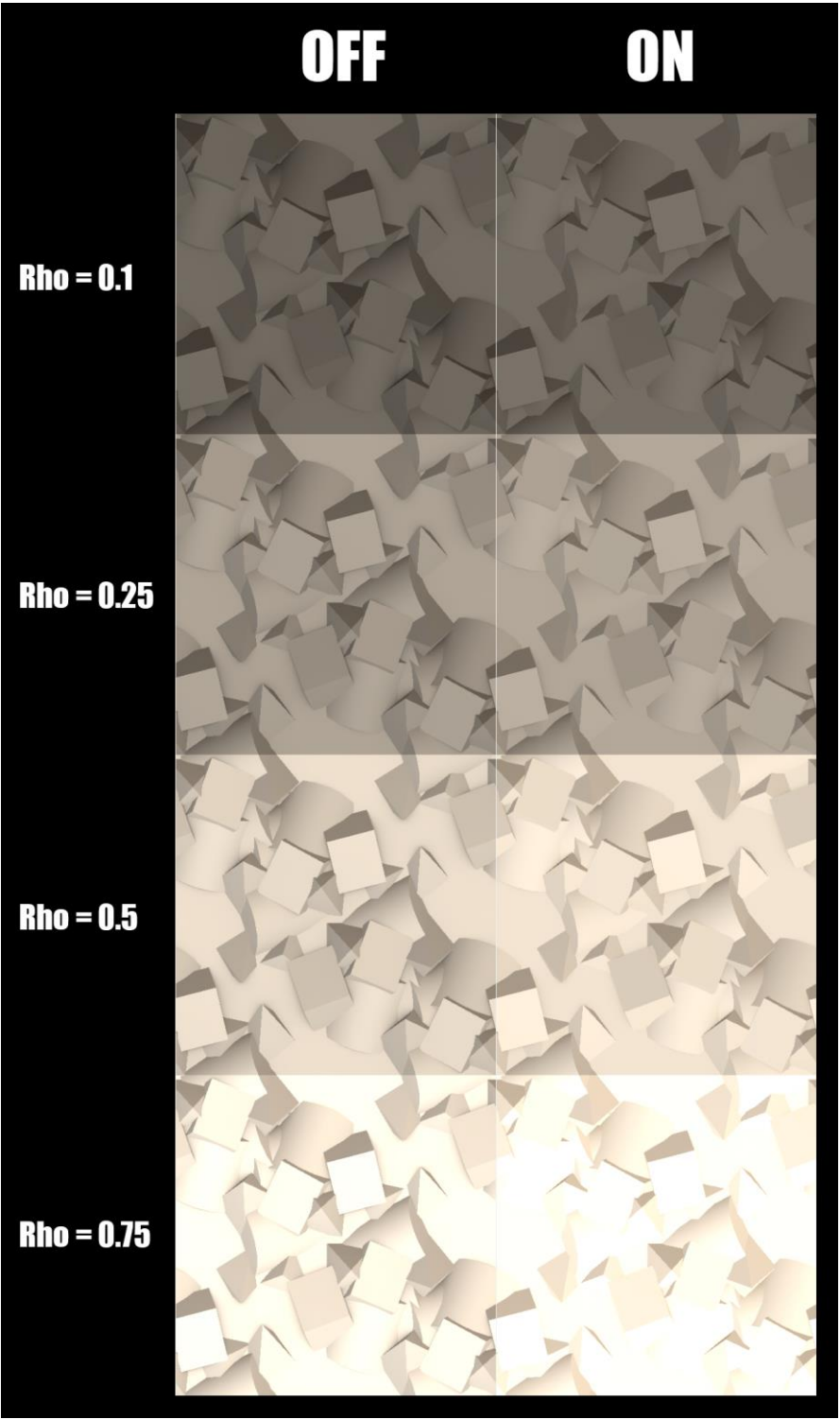
RESULTING CURVE FITTING FOR VARIOUS VALUES OF ρ FROM 0 TO 100% EVERY 10%.
RED: EXPERIMENTAL DATA. THICK CURVES: FITTED MODEL.

We should pay special attention to the $\frac{\rho(x)}{1-\rho(x).\tau}$ 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:

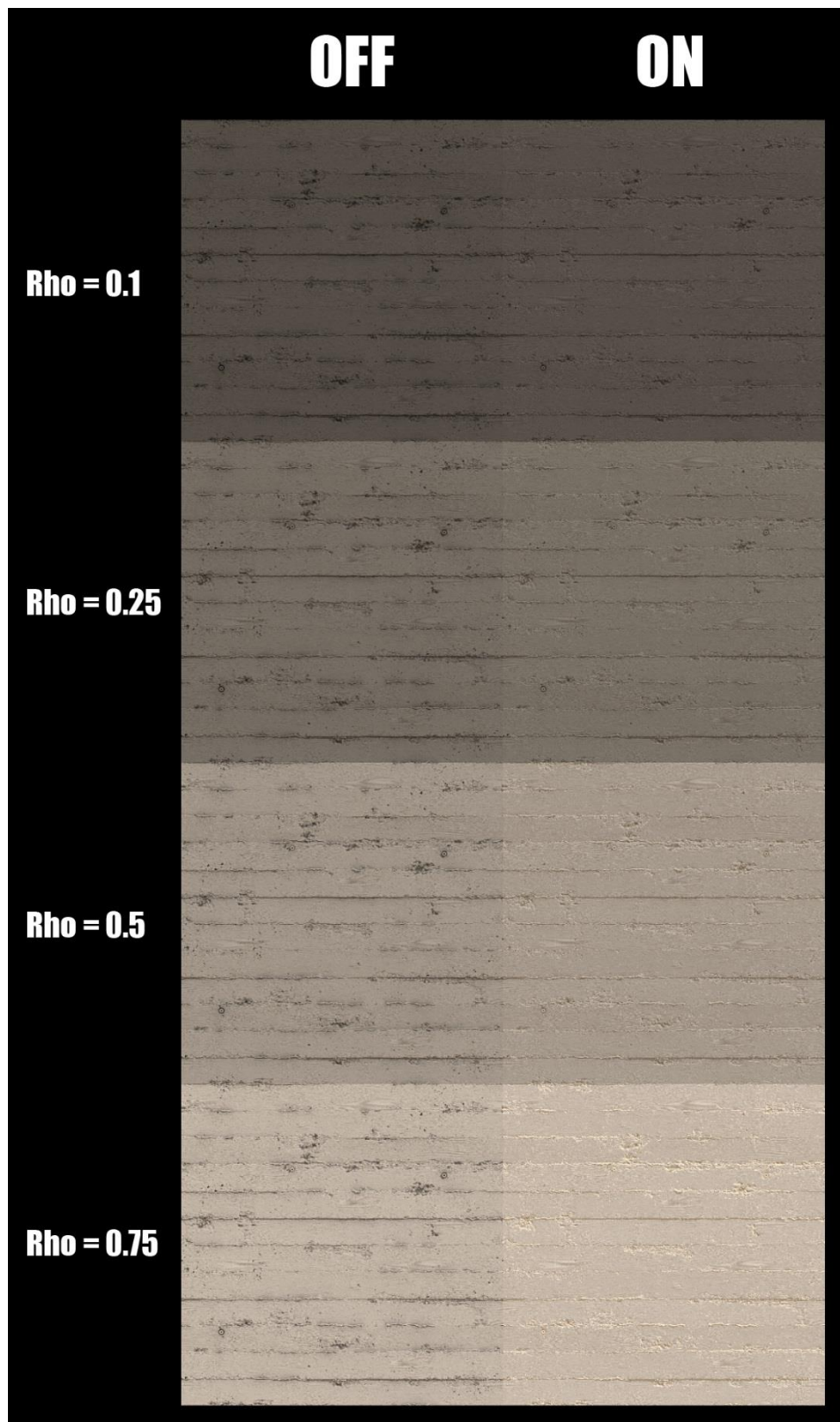


THE EFFECT OF ALBEDO SATURATION FOR VARYING VALUES OF τ FROM 0 TO 1

Showtime



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



Concrete – 1m x 1m, 5cm Elevation

Improving even Further : Bent Normal and Cone Aperture

#TODO!!! ➡ Modifier le soft de génération d'AO pour pisser un « Bent Cone » (=Bent Normal + Cone d'Ouverture) + Montrer l'utilisation et notamment la modification de l'intégrale d'irradiance en SH (reprendre le calcul de ma page wiki).