Microfacet Model Regularization for Robust Light Transport

Johannes Jendersie and Thorsten Grosch,
July 10

Notes

Hello everybody and wellcome to my talk about microfacet-based regularization.

Challenge: Complex Light Paths

Path Tracing 1000 spp — Reference (VCM) 155k spp

- Path Tracing (PT) is widely used because of its simplicity
- Even more complex algorithms like Vertex Connection and Merging (VCM) cannot handle highly glossy paths well

Notes

The challenge to solve is that of hard to sample light paths. In modern scenes many different materials are used. For example, the specular-diffuse-specular paths, which we see in the dial of the watch, lead to problems in many renderers. Beyond that paths with only glossy and specular materials will still have a high variance in expensive methods like vertex connection and merging. Even after as many as 155000 samples we can still see noise in the vcm renderering.
Thus, our goal is to reduce the variance in any method – especially in the cheaper methods like path tracing.

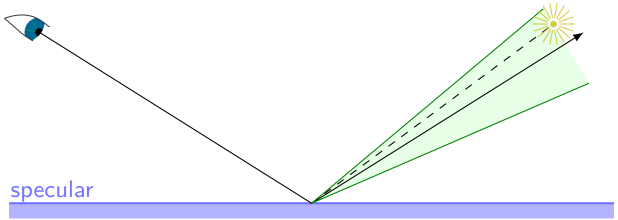
BSDF Regularization as Solution

- Sharp peaks in the *Bidirectional Scattering Distribution Function* (BSDF) of a material cause the problems
- Blurring the BSDF should help

Notes

So, sharp peaks in the *Bidirectional Scattering Distribution Function* of a material cause the problems. Therefore, smoothing the BSDF should help.
And it does as these two images show. Both are rendered with path tracing using 1000 spp where the right one uses the regularization strategy which I will explain in the following.

Regularization of Specular Events



Path Space Regularization for Holistic and Robust Light Transport
Kaplanyan and Dachsbacher [KD13]

Improving Robustness of Monte-Carlo Global Illumination with Directional Regularization
Bouchard et al. [BIOP13]

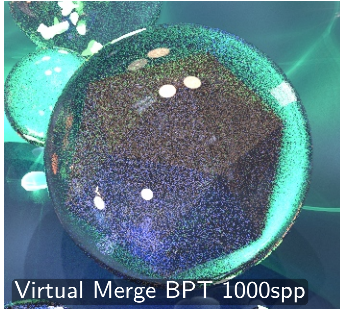
Notes

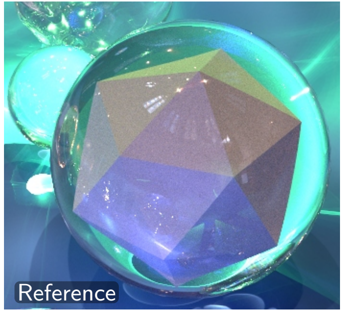
Kaplanyan and Dachsbacher introduced the concept of light path regularization to handle specular events.
The idea is to accept any connection which is within a small cone around the actual reflection direction.
Bouchard et al. used regularized and non-regularized samplers at the same time and let a MIS heuristics decide which to use.

Regularization of Glossy Events

Virtual merge strategy:

- If connecting, create a random sample of the current BSDF
- Check if BSDF within cone
  - Discard if not
  - If yes, multiply the sample value with the solid angle of the cone

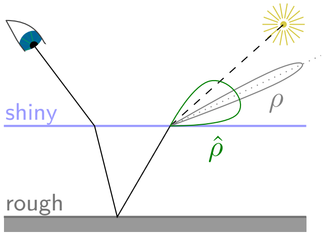




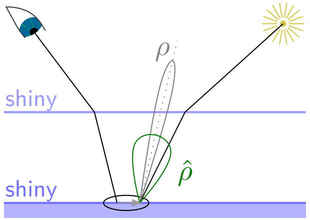
Notes

The same idea can be used for arbitrary reflectance models which I call the virtual merge strategy. We can think of it as a sampling event which is accepted iff the random sample comes close enough to the target point of the connection.
While having sharp, but enlarged highligths this method has a high variance. The reason is the additional random sampling we need.
To avoid the random decision we would need to compute a closed form integral of the BSDF over the cone. This however, is not possible for many models.

Regularization of Glossy Events



In connections



In merges

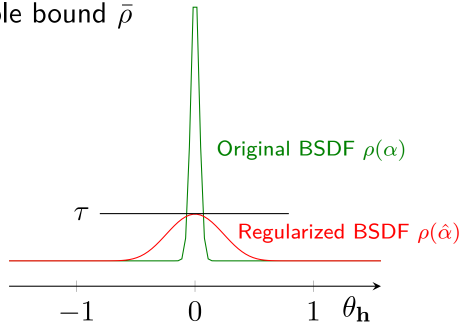
- Changing roughness α of a microfacet model in connections and merges
- Non-regularized BSDF during random walk!

Notes

Alternatively, we can control the roughness in microfacet models, which is often parametrized by α.
If we do, a connection or a merge on a smooth surface can have non-zero contributions around the real scattering directions. Here, ρ is the BSDF as it should be and $\hat{\rho}$ is the regularized variant with an increased roughness.
Note, that we do not change the BSDFs in random walks. Doing so would only increase bias and variance as I will explain later.

How to Control the Amount of Smoothness?

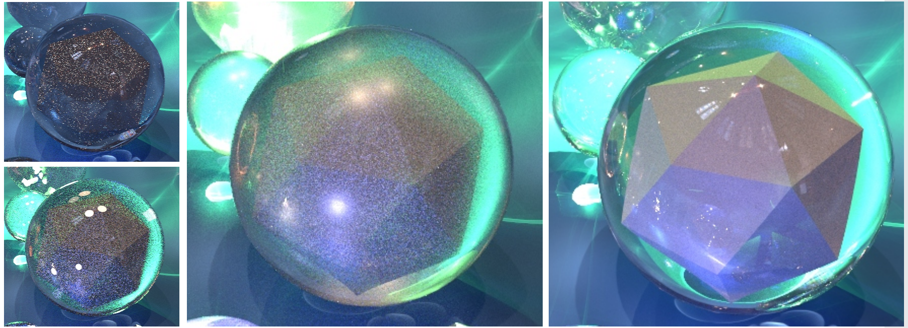
- We want to invert $\max(\rho(\hat{\alpha})) = \tau$ for $\hat{\alpha}$
- Closed form not possible in general
- Instead we search an invertible bound $\bar{\rho}$
  - Basically $\bar{\rho} \propto {}^{1}/_{\pi\alpha^2}$
  - Bound is not strict



Notes

To control the amount of smoothness, I introduced a threshold prameter τ for the maximum value of a BSDF. Since the maximum value of a normalized function has a proportional influence on the upper limit of variance, this is a logic choice.
So, we search the parameter α for which tau will be the maximum value of the BSDF. This requires an inversion of the BSDF which is in general not possible in a closed form. Since we cannot invert the BSDF directly we use an invertible bound instead. The bound derived in the paper is basically proportional to 1 by pi alpha square.
However, the bound is not strict when using the v-cavity shadow model. At grazing angles the BSDF always becomes infinite independent of alpha.

Comparison: Bidirectional PT 1000spp



Top: No regularization
Bottom: Virtual Merge

Roughness-based reg.

Reference

Notes

Here, we see the results of roughness based regularization in a bidirectional path tracer. The variance is lower than it is without regularization or with the virtual merge strategy. Instead, highlights get smoothed a lot. This is a problem which we will face later in this presentation.

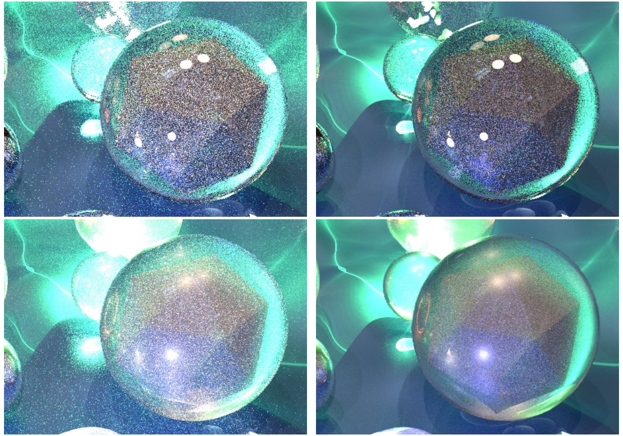
Corrected MIS Weights

Notes

The next problem we need to solve is that the MIS weights must reflect the change we made to the BSDF.

Correct MIS Weights

Independent of the method an unmodified MIS weight will fail



Unmodified (invalid) MIS   Correct MIS

Notes

Independent of the regularizaton method the unmodified MIS from any rendering algorithm will fail, as we see here on the left. The result is not only more noisy, but it is also too bright. On the right we can see results using the same sample set with a corrected weighting scheme.

Correct MIS Weights

- The probability to find a path did not change!
- Our approach changes the measurement contribution function f

$$f = \dots G(\mathbf{x}_{i-1}, \mathbf{x}_i)\rho(\mathbf{x}_i, \alpha_i)G(\mathbf{x}_i, \mathbf{x}_{i+1}) \dots$$

- Each connection or merge changes f at a different point i by exchanging a single α
- Reevaluated the derivation of the balance heuristic [Vea97]

$$\Rightarrow w_i = \frac{n_i(1/\hat{I}_i)}{\sum_k n_k(1/\hat{I}_k)} \quad \text{not} \quad \frac{n_i p_i}{\sum_k n_k p_k}$$

- Implementation possible through a small modification of p

Notes

First note that we did not change the sampling of the path and therefore the path probability did not change. So, why do we need to change the MIS?
Our approach changes the measurement contribution function small f. It consists of a number of geometric transport terms and the BSDFs. Each connection or merge changes f at a different position $\mathbf{x}_i$ by exchanging a single alpha. This means that f differs between samplers and is not constant anymore.
We reevaluated the derivation of the balance heuristic which shows that we need to use the sample values $\hat{I}_i$ and $\hat{I}_k$ instead the probabilities. However, it is equivalent to modify the probabilities to minimize the implementation impact on existing renderers. For details please have a look in the paper.

Variance and Bias

Notes

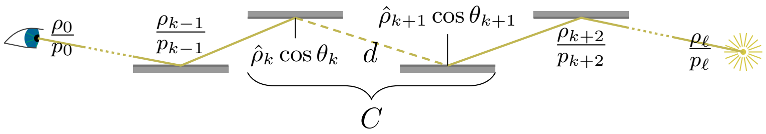
Having a working basic implementation we can focus on the error which will arise.

Variance Considerations

How much does BSDF regularization reduce the variance?

- Each sampler consists of a series of sampling events ρ/p and a central connection or merge term C

$$V\left[\prod_i \frac{\rho_i}{p_i} \cdot C\right]$$



Notes

The question I want to answer next is: How much does BSDF regularization reduce the variance.
Each sampler consists of a series of sampling events and some other terms which are summarized in C.

Variance Considerations

- Product can be decomposed
  $V[XY] = E[X]^2 V[Y] + E[Y]^2 V[X] + V[X]V[Y]$
- Term for sampling part

$$V\left[\prod_i \frac{\rho_i}{p_i}\right] \approx 0 \qquad E\left[\prod_i \frac{\rho_i}{p_i}\right] = \prod_i c_i$$

- Most importance sampler have a very good fit to the BSDF
  - Lambert: perfect $V = 0$
  - Specular: perfect $V = 0$
  - Microfacet: very good [Hd14]
- Variance is not directly caused in the sampling part
- Clamping or smoothing ρ/p will not reduce variance

Notes

The product variance can be decomposed using a basic property of the variance. First, we observe that only little variance comes from the sampling itself.
Most importance sampler have a very good fit. We are able to perfectly sample Lambertian diffuse and specular surfaces without variance in this part of the estimator. Also, there is a sufficient good sampler for the microfacet models which was invented by Heitz and d'Eon.
Variance is not directly caused in the sampling part. Therefore, clamping or smoothing the sampling throughputs will not reduce the variance unless the material sampling routine is bad.

Variance Considerations

- Variance originates in the remaining central term C

$$V\left[\frac{\hat{\rho}_k \cos\theta_k \cos\theta_{k+1}\hat{\rho}_{k+1}}{d^2}\right] \gg 0$$

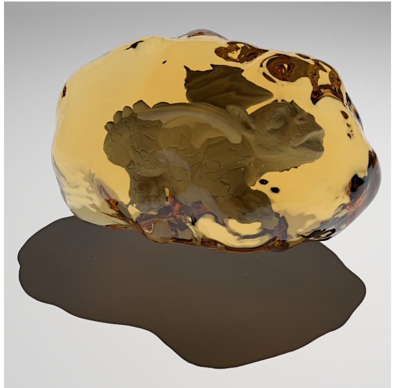
$$\text{and } V\left[\frac{\hat{\rho}_k}{\pi r^2}\right] \gg 0.$$

- The variance is large, if the terms vary greatly over the footprints of the two sub-paths
- Footprint = distribution of sub-path end points
  $\Rightarrow$ Strong scattering makes a high variance likelier
- Roughing $\rho \to \hat{\rho}$ decreases the variation of the terms
- All variance beyond the variation of ρ cannot be removed through regularization – regardless of the method

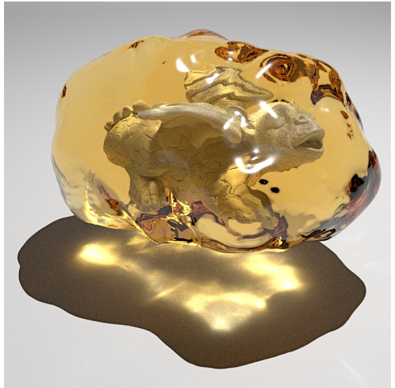
Notes

The point where the variance originates is in the remaining terms of connections and merges. The variance is large, if the terms vary greatly over the footprints of the two sub-paths
That is, if a direction is scattered completely random, but any point which is hit afterwards has the same brightness, there is still no variance. Hence, diffuse sampling does not introduce variance by itself, but it increases the footprint of the path. The larger the spread of this end points the liklier it is that the above terms will introduce variance.
Now, smoothing the BSDF decreases the variation of the central terms and consequently reduce the variance for any possible footprint distribution.
However, all variance beyond that, for example from the cosine or the distance terms or from sampling itself, cannot be removed through regularization. This is the case for any BSDF-based approach, not only our proposed solution.

Intermediate Results



PT 16k spp



Regularized PT 16k spp

Notes

Let us look at some first results. Here we see a path tracer with 16000 sampler per pixel. In the non-regularized image on the left there are neither caustics nor SDS paths. The right image looks good, but has over-blurry highlights which could be avoided.

Bias Reduction: Path Diffusion

- Recap: Large footprint $\Rightarrow$ larger chance for variance
- Almost deterministic paths $\Rightarrow$ less blurring required
- Modify τ with the tangential standard deviation at vertex k

$$\tau' = \frac{\tau}{\sigma_{\triangleleft,k}}$$

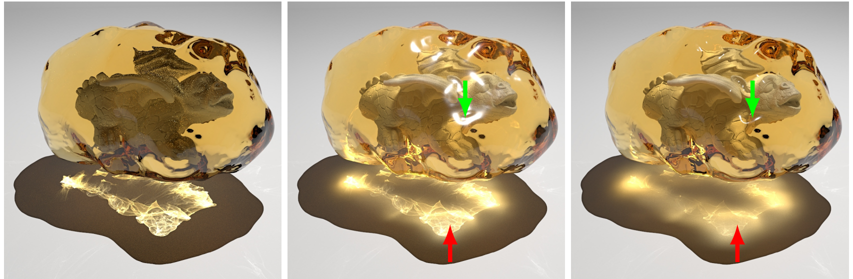
- Simplified 5D covariance tracing [BSS*13] to compute $\sigma_{\triangleleft}$.

Notes

Which brings us to the first bias reduction heuristic.
From the previous variance analysis we know that large footprints cause larger variances. On the other hand, almost deterministic paths have a small footprint and thus less blurring is required. In these situations we can increase the threshold to decrease the bias of regularization.
As a solution we modify the threashold by dividing it with the tangential standard deviation of incoming directions at vertex k. We calculate this standard deviation with a simplified form of the 5D covariance tracing by Belcour et al.

Bias Reduction: Path Diffusion in BPT

BPT with 16k spp without regularization (left), with regularization (center) and with path diffusion heuristic (right)



- Success: the highlight (and other near specular paths) are sharper
- Problem: Caustic is blurred more (correctly assumes that the connection to the light has high variance ⇒ strong regularization)

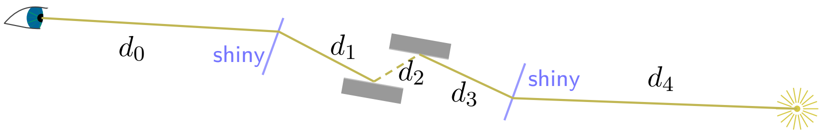
Notes

Let us see how it works in practice. This time we have a bidirectional path tracer which is capable of rendering caustics.
Enabling regluarization will add the missing SDS path, but also blurs the caustics and highlights.
The path diffusion heuristic successfully reduces the blurryness of highlights.
Unfortunately, it also increases the blurryness of the caustic, because after the diffuse reflection on the ground the next events will have a high angular standard deviation. That means, the diffusion heuristic correctly assumes that, for example the connection to the light, requires strong regularization to reduce the variance. However, it would not have been necessary to regularize any of the samplers.

Bias Reduction: Sampler Quality

- If there is a good sampler, other samplers do not need to be regularized at all
- Sampler Quality

$$q = \min_k \left(\underbrace{\bar{\rho}_k \cdot \bar{\rho}_{k+1}}_{\text{Bound on connection term}} \cdot \left({}^{d_{\text{path}}}/_{d_k}\right)^2 \right) \qquad d_{\text{path}} = \sum d_i$$

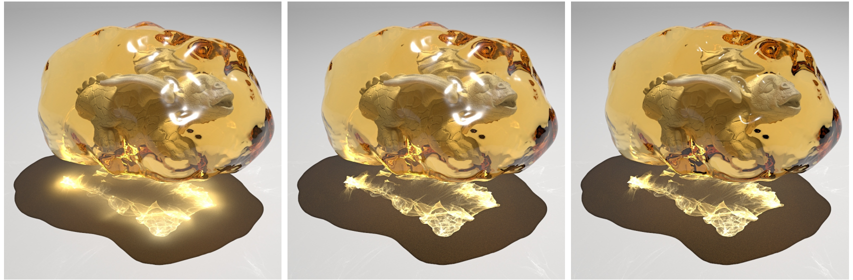


Notes

So, we introduce another heuristic – the sampler quality. If there is a good sampler we can disable regularization for all other samplers on the same path.
To assess the quality of the best sampler on the path we take the minimum over the bound terms of connections.
And we multiply with the squared relative path length to get a measure of the shortness of a connection.

Bias Reduction: Sampler Quality in BPT

$$\tau' = \begin{cases} \tau & \text{if } q \geq \tau^2 \\ \infty & \text{otherwise} \end{cases}$$



Non-adaptive | Sampler quality | Path Diffusion Sampler quality

Notes

If there is a sampler with a high quality, which has a small value q, we set the threshold to infinity to disable regularization. Since the bound on the previous slide contains two BSDF bounds it makes sense to compare q to the squared threshold which we apply on a per BSDF basis.
In the images we can see that the sampler quality heuristic successfully preserves the caustic. On the right we can also see that a combination of both heuristics reduces bias quite successfully.

Results and Outlook

Notes

Let us come to a conclusion.

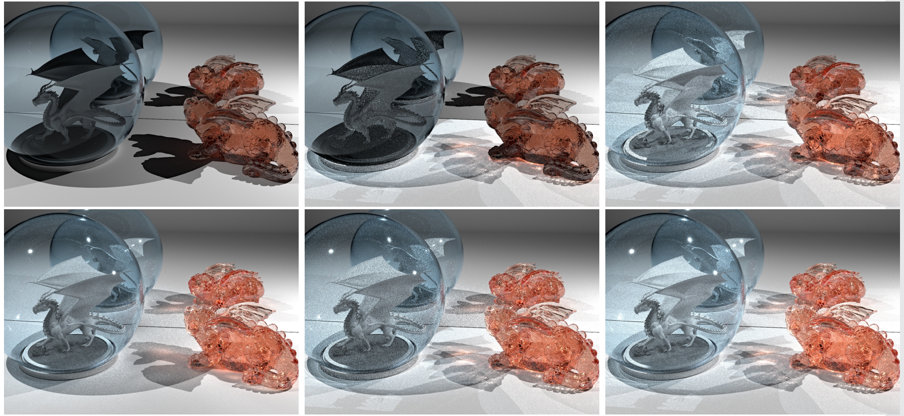
Results

- BSDF regularization can turn (almost) impossible paths into unlikely ones
- Variance can be reduced but never guaranteed
  - Reason: other terms stay unchanged
- MIS weight must be adapted, because regularization changes f
- Heuristics help to avoid bias
- Opposed to clamping, our approach keeps most of the energy
  - Microfacet models lose energy due to missing multiple scattering

Notes

BSDF regularization can turn (almost) impossible paths into unlikely ones.
Variance can be reduced but never guaranteed, because other terms which contribute variance stay unchanged.
And finally, we can reduce the most of the unnecessary bias with two heuristics.
Opposed to clamping, increasing roughness keeps most of the energy.
Why most? Well, if not explicitly compensated, microfacet models lose energy due to missing multiple scattering.

Equal time 1h

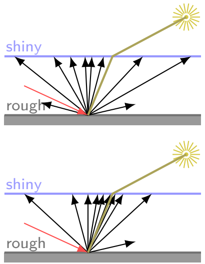


PT (1015 / 1008 spp) BPT (541 / 523 spp) VCM (487 / 493 spp)

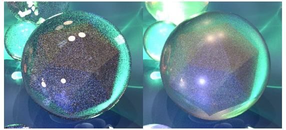
Notes

Here, we see regularzation applied in different rendering algorithms. As shown before, path tracing and bidirectional path tracing miss some important effects. Indeed, all three regularized methods lead to very similar results. Only, the caustics from regularization are more blurry than that without and also contain more noise.
This is because regularization only turns almost impossible into unlikely paths, while other sampling methods, like merges, can find the same paths very well.

Future Work



- Guided sampling ⇒ More samples close to the contributing direction
  - *Product Importance Sampling* [HEV*16]
  - *Practical Path Guiding* [MGN17]
- Improved Connection Samplers ⇒ Promising variance reduction
  - *Probabilistic Connections for Bidirectional Path Tracing* [PRDD15]
  - *Matrix Bidirectional Path Tracing* [CBH*18]
- Control Variates ⇒ better virtual merges



Questions?

Notes

Bouchard G., Iehl J.-C., Ostromoukhov V., Poulin P.:
Improving Robustness of Monte-Carlo Global Illumination with Directional Regularization.
In *SIGGRAPH Asia 2013 Technical Briefs* (2013), SA '13, ACM, pp. 22:1–22:4.
URL: http://doi.acm.org/10.1145/2542355.2542383.

Belcour L., Soler C., Subr K., Holzschuch N., Durand F.:
5d Covariance Tracing for Efficient Defocus and Motion Blur.
ACM Transaction on Graphics (TOG) 32, 3 (July 2013), 31:1–31:18.
URL: http://doi.acm.org/10.1145/2487228.2487239.

Chaitanya C. R. A., Belcour L., Hachisuka T., Premoze S., Pantaleoni J., Nowrouzezahrai D.:
Matrix Bidirectional Path Tracing.
In *Proc. of Eurographics Symposium on Rendering (EGSR)* (2018), pp. 023–032.
doi:10.2312/sre.20181169.

Heitz E., d'Eon E.:
Importance Sampling Microfacet-Based BSDFs Using the Distribution of Visible Normals.
Computer Graphics Forum (CGF) 33, 4 (2014), 103–112.
URL: http://dx.doi.org/10.1111/cgf.12417.

Heitz E., Dupuy J., Hill S., Neubelt D.:

Notes

Real-time Polygonal-light Shading with Linearly Transformed Cosines.
ACM Transaction on Graphics (TOG) 35, 4 (July 2016), 41:1–41:8.
URL: http://doi.acm.org/10.1145/2897824.2925895.

Herholz S., Elek O., Vorba J., Lensch H., Křivánek J.:
Product Importance Sampling for Light Transport Path Guiding.
Computer Graphics Forum (CGF) 35, 4 (2016), 67–77.
URL: http://dx.doi.org/10.1111/cgf.12950.

Kaplanyan A., Dachsbacher C.:
Path Space Regularization for Holistic and Robust Light Transport.
Computer Graphics Forum (CGF) 32 (2013), 63–72.
URL: http://dx.doi.org/10.1111/cgf.12026.

Müller T., Gross M., Novák J.:
Practical Path Guiding for Efficient Light-Transport Simulation.
In *Proc. of Eurographics Symposium on Rendering (EGSR)* (June 2017).
URL: http://onlinelibrary.wiley.com/doi/10.1111/cgf.13227/full.

Popov S., Ramamoorthi R., Durand F., Drettakis G.:
Probabilistic Connections for Bidirectional Path Tracing.
Computer Graphics Forum (Proc. of EGSR) 34, 4 (2015), 75–86.
URL: http://dx.doi.org/10.1111/cgf.12680.

Torrance K. E., Sparrow E. M.:

Notes

Theory for Off-Specular Reflection From Roughened Surfaces.
Journal of the Optical Society of America 57, 9 (Sept. 1967), 1105–1114.
URL: https://doi.org/10.1364/JOSA.57.001105.

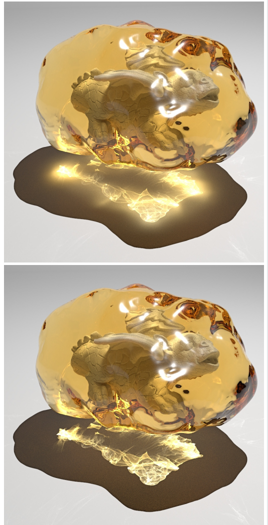
Veach E.:
Robust Monte Carlo Methods for Light Transport Simulation.
PhD thesis, Stanford University, 1997.
URL: http://graphics.stanford.edu/papers/veach_thesis/.

Walter B., Marschner S. R., Li H., Torrance K. E.:
Microfacet Models for Refraction Through Rough Surfaces.
In *Proc. of Eurographics Symposium on Rendering (EGSR)* (2007), Eurographics Association, pp. 195–206.
URL: http://dx.doi.org/10.2312/EGWR/EGSR07/195-206.

Notes

Problem: Discontinuity of Decisions

- A small change of τ or the path length may change the decision
- Problem is intrinsic
  - Let $\tau^2 = 100$ and $q = 100.1$
  - Assume the best sampler (with $q = 100.1$) cannot be regularized
  - All other samplers have $q_k \gg q$
  - At least on other sampler must be regularized extremely to be $\leq \tau^2$
- Smoothstep can reduce the problem, but an arbitrary slope must be set



Notes

Unfortunately, the sampler quality heuristic can cause temporal incoherent decisions. A small change of the threshold or the path length can toggle the decision.
This problem is not only caused by the step function. Rather, it is intrinsic to the problem of guaranteing one good sampler. For example let τ^2 be 100 and the best sampler have $q = 100.1$. Now, this best sampler has two Lambertian diffuse vertices and cannot be regularized. So, we must regularize one of the other samplers which have much more variance in the beginning. Thus, in the moment where the threshold is failed by a tiny amount we introduce a huge bias in one sampler to compensate that. Especially, this is not a problem of our week variance estimation. Even if we would know the exact variance of each sampler the situation stays the same. To guarantee a low variance sampler with the tool of BSDF regularization will always cause this discontinuity.
We experimented with different smoothstep functions, but those always failed in one or the other situation so we used the simplest possible function here. The problem with smoothstep functions is that one must define an additional parameter to stir the rate of change. A flat transition will cause blurred caustics and a steep one will have more temporal inconsistency.

True Variance?

- Double integral over the two footprints
- Approximate: Footprint area times pdf bound
- Bhatia-Davis bound on pdfs: $V \leq (M - \mu)(\mu - m)$
- Weaker Popoviciu bound: $V < \frac{1}{4}(M - m)$

Notes

Computing α

- We want to invert $\rho = \tau$ for α
- Closed form not possible in general
- Instead we search an invertible bound $\bar{\rho}$

Reflection [TS67] $$\bar{\rho}_r(\alpha) = \frac{1}{4\pi\alpha^2}\bar{G}(\alpha)$$

Refraction [WMLT07] $$\bar{\rho}_t(\alpha) = \frac{\max(\eta_i, \eta_t)^2}{\pi\alpha^2(\eta_i - \eta_t)^2}\bar{G}(\alpha)$$

with $$\bar{G}(\alpha) = \begin{cases} 1 & \text{V-cavity} \\ 4/\alpha^2 & \text{Smith, GGX} \\ 4\pi/\alpha^2 & \text{Smith, Beckmann or Cosine} \end{cases}$$

- Bound is not strict for V-cavity

Notes

This search for alpha requires an inversion of the BSDF which is in general not possible in a closed form. Since we cannot invert the BSDF directly we use an invertible bound instead. The bound derived in the paper looks like this.
However, the bound is not strict when using the v-cavity shadow model. At grazing angles the BSDF always becomes infinite independent of alpha.

MIS Weights for Estimators

- Original derivation by Veach [Vea97]:

$$F = \sum_{i=1}^{m} \frac{1}{n_i} \sum_{j=1}^{n_i} w_i(\mathbf{x}_j)\hat{I}_i(\mathbf{x}_j)$$

$$V[F] = \left(\int_{\Omega} \sum_{i=1}^{m} \frac{w_i^2(\mathbf{x})}{n_i} \hat{I}_i(\mathbf{x})^2 p_i^*(\mathbf{x})\ \mathrm{d}\mu(\mathbf{x})\right) - \left(\sum_{i=1}^{m} \frac{1}{n_i} \mu_i^2\right)$$

$$\Rightarrow w_i = \frac{n_i(1/\hat{I}_i)}{\sum_k n_k(1/\hat{I}_k)} \quad \text{not} \quad \frac{n_i p_i^*}{\sum_k n_k p_k^*}$$

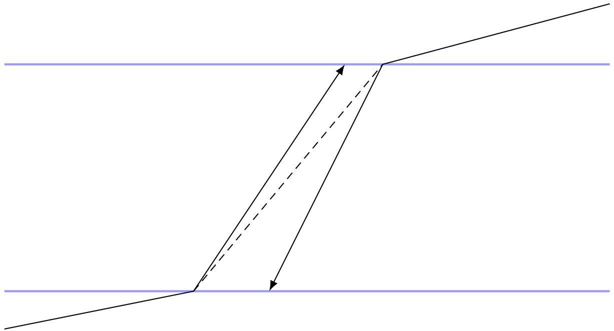
- Common simplification: f is the same for all estimators $\hat{I}_i = f/p_i^*$

$\Rightarrow$ can be canceled

Notes

To find the correct solution we restart with the derivation of the balance heuristic. We estimate a Monte Carlo integral F over samples I_i with weights w. As shown in Veach's PHD thesis F has a variance which looks this term. By minimizing the left of the two terms through optimizing w one obtains the balance heuristic using the sample values directly and not the more known form.
Usually, this can be simplified, because each sample has the form: path measurement function f divided by the path probability p^*. Therefore, f can be canceled out and we get the form on the right side. However, for regularization f changes for different estimators and cannot be removed.

Why Regularization with Standard MIS is Biased



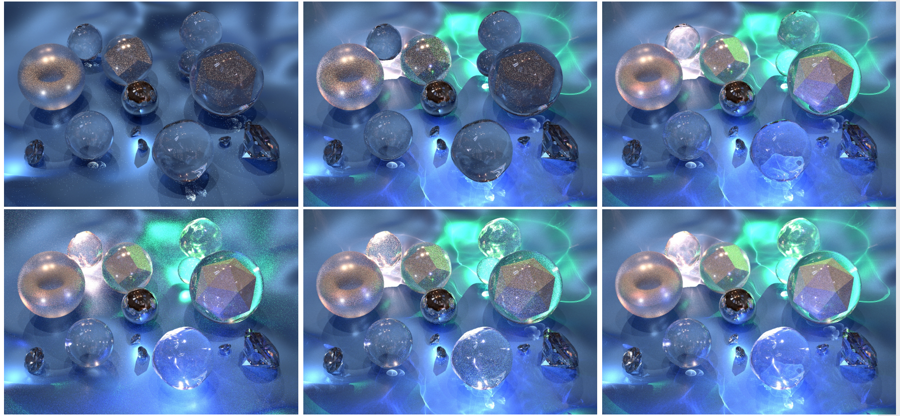
- Two specular surfaces $\Rightarrow \rho = p = 0$ but $\hat{\rho} > 0$
- MIS weight uses $p = 0$ for all other samplers $\Rightarrow w = 1$

$\Rightarrow$ Each sampler does the same $\Rightarrow \forall w_i = 1$ which leads to an overestimation by a factor of the number of path segments

- Same is true to a lesser degree for glossy events

Notes

Equal time 1h



PT (970 / 941 spp) BPT (647 / 625 spp) VCM (432 / 433 spp)

Notes

Here is yet another scene with complex light paths. Again, all three regularized variants produce similar but different noisy results. Summarizing, I would prefer the more expensive sampling algorithms, where regularization still helps to find near specular paths. This increases the fidelity of specular objects like this glass sphere or the dragon on the previous slide.