Gradient-Domain Path Tracing

Pramook Khungurn

September 26, 2015

This document is written as I read the "Gradient-Domain Path Tracing" paper [Kettunen et al., 2015].

1 Introduction

- There has been recent works that extends Metropolis light transport to compute gradients of the resulting image instead of computing the pixel colors.

- This paper shows that you can do that in normal path tracing as well.

  - Instead of tracing a single path, shoot an additional finite-difference path shifted by one pixel. This gives you image contribution of each path and also an estimate of the finite difference between the two pixels.

  - In the end, perform a screened Poisson reconstruction that combines all information: both the image contributions and the gradient estimate.

- The gradient is not estimated from difference of two uncorrelated random paths. The paths generated are highly correlated, and this lowers the variance.

- Gradient rendering yields better quality than direct value rendering when the resuling image has less energy in high frequencies than in low frequencies. The screend Poisson reconstruction can get the best of both worlds.

- The paper introduce new sampling strategies:

  - A new shift mapping whose Jacobian is very easy to calculate.

  - The gradient can either be approximated with forward or inverse shift. These can be combined with multiple importance sampling to yield better results.

- As a note, the screen Poisson reconstruction takes as input the gradient image $I^{\mathrm{d}x}$ and $I^{\mathrm{d}y}$ and a rough estimate of the actual image I^g. It solves for the image I by solving the following minimization problem:

$$\arg\min_{I} \left(\left\| \begin{pmatrix} H^{\mathrm{d}x}I \\ H^{\mathrm{d}y}I \end{pmatrix} - \begin{pmatrix} I^{\mathrm{d}x} \\ I^{\mathrm{d}y} \end{pmatrix} \right\|_2^2 + \|\alpha(I - I^g)\|_2^2 \right)$$

where $H^{\mathrm{d}x}$, $H^{\mathrm{d}y}$ are matrices that take finite differences along both rows and columns and returns appropriately vectorized gradients, and α is a weighting parameter.

2 Overview

- Image gradient = differences between pairs of neighbor pixels.

- Gradient-domain rendering samples image gradients instead of pixel values.

- In gradient-domain path tracing, each *base path* is paired with an *offset path* through the neighboring pixel. See the picture below for what they look like.
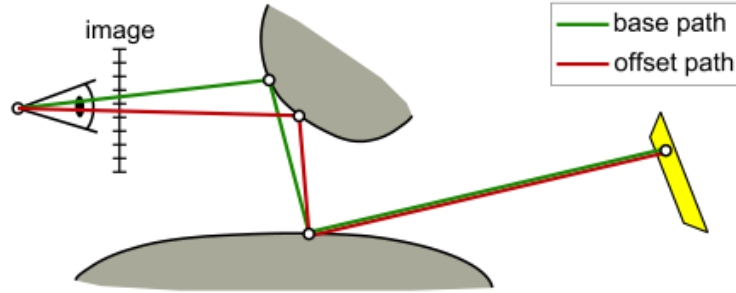
- The *shift mapping* computes the offset path from the base path.

- The difference between the contribution of the two paths gives an estimate of the gradient.

- More precisely, let $\Delta_{i,j}$ denote the difference between pixels i's and j's contribution:

$$\Delta_{i,j} = \left(h(x) * \int_\Omega f(x, \bar{p}) \, \mathrm{d}\mu(\bar{p}) \right)(x_i) - \left(h(x) * \int_\Omega f(x, \bar{p}) \, \mathrm{d}\mu(\bar{p}) \right)(x_j)$$

where:

- x is the image coordinate,
- $h(x)$ is the pixel filter,
- Ω is the space of all light paths,
- $(x, \bar{p})$ is the path where $\bar{p}$ denotes all the path configurations from a point on the sensor to a light source,
- f is the image contribution function.

The pixel filters are evaluated at pixel centers x_i and x_j.

- The above definition of $\Delta_{i,j}$ are what was used in previous works on gradient-domain rendering. The paper, however, tries to estimate the difference with a single integral:

$$\Delta_{i,j} = \left(h(x) * \int_\Omega f(x, \bar{p}) - f(T_{ij}(x, \bar{p}))|T'_{ij}| \, d\mu(\bar{p}) \right) = \left(h(x) * \int_\Omega g_{ij}(x, \bar{p}) \, d\mu(\bar{p}) \right)$$

where $T_i j$ is the shift mapping that deterministically maps a base path $(x, \bar{p})$ to an offset path $T_{ij}(x, \bar{p})$ such that they are close to each other in path space. (The difference is only at the intersection?)

- The T_{ij} function only allows shifting by one pixel distance so that the filter value at the offset path is the same as that of the base path. In this way, a single convolution is enough to define the gradient.

3 Theoretical Analysis

3.1 Error Analysis of Gradient Estimate

- To enable Fourier analysis, the paper makes the following assumptions:

  - The image is 1D.
  - Paths are parameterized over a Cartesian hypercube, like the one used in Kelemen-style Metropolis light transport.
  - The first path dimension is the image axis.
  - All terms related to measure in path space are hided.
  - Analysis is only restricted to uniform sampling in the path space parameterization. (No importance sampling is allowed?)
  - Sampling is statinary stochastic processes. Basically, it means that it extends over an infinite domain, which implies that there is no boundaries to worry about. This, however, tends to overestimate the variance of the real case in which the domain is finite.

- The image contribution function f is defined over an image axis x and an orbitrary long vector of path parameters $\bar{p}$ (assumed to be the Cartesian hypercube like in Kelemen-style Metropolis light transport).

- The shift mapping T is given by $T(x, \bar{p}) = T(x - 1, \bar{p})$, which has unit gradient.

- The path difference function g is simply $g(x, \bar{p}) = f(x, \bar{p}) - f(x - 1, \bar{p})$, which can be written as the convolution $g(x, \bar{p}) = (d * f)(x, \bar{p})$ where $d(x, \bar{p}) = \delta(x) - \delta(x - 1)$.

- In Fourier domain, $g = d * f$ becomes $G = DF$. The power spectrum of the path different function is given by:

$$|G(\omega_x, \omega_{\bar{p}})|^2 = |D(\omega_x, \omega_{\bar{p}})|^2 |F(\omega_x, \omega_{\bar{p}})|^2$$
$$= (2 - 2cos(2\pi\omega_x))|F(\omega_x, \omega_{\bar{p}})|^2$$

where ω_x and ω_p are frequencies over the image and path space.

- We can observe from the above equation some well-known results:

  - Finite difference cancels the DC. This is because $|D(0, \omega_{\bar{p}})|^2 = 0$.
  - Finite difference attenuates low frequencies because $|D(\omega_x, \omega_{\bar{p}})|^2$ is low when ω_x is low.
  - Finite difference boosts square magnitudes of high frequencies by a factor up to four for the Nyquist limit $\omega_x = 1/2$ of the image. This is because $|D(1/2, \omega_{\bar{p}})|^2 = 4$.

- Given that the sampling gives rise to unbiased estimates, the variance is characterized by the mean square error (MSE).

- Under a number of assumption, the expected squared errors in frequency domain $|\epsilon_F(\omega_x, \omega_{\bar{p}})|^2$ and $|\epsilon_G(\omega_x, \omega_{\bar{p}})|^2$ for all frequencies are constant:

$$|\epsilon_F(\omega_x, \omega_{\bar{p}})|^2 = \frac{1}{n}\|F\|^2$$
$$|\epsilon_G(\omega_x, \omega_{\bar{p}})|^2 = \frac{1}{n}\|G\|^2$$

3

where

$$\|F\|^2 = \int |F(\omega_x, \omega_{\bar{p}})|^2 \ \mathrm{d}\omega_x \mathrm{d}\omega_{\bar{p}}$$

$$\|G\|^2 = \int |G(\omega_x, \omega_{\bar{p}})|^2 \ \mathrm{d}\omega_x \mathrm{d}\omega_{\bar{p}} = \int (2 - 2\cos(2\pi\omega_x))|F(\omega_x, \omega_{\bar{p}})|^2 \ \mathrm{d}\omega_x \mathrm{d}\omega_{\bar{p}}$$

- Further analysis gives the MSE in the image domain only (by slicing the frequency domain at $\omega_{\bar{p}} = 0$). Assuming perfect reconstruction kernel h, we have:

$$|\epsilon_F(\omega_x)|^2 = |\epsilon_F(\omega_x, 0)|^2 = \frac{1}{n}\|F\|^2, \ \text{if } |\omega_x| < 1/2, \text{ otherwise } 0,$$

$$|\epsilon_G(\omega_x)|^2 = |\epsilon_G(\omega_x, 0)|^2 = \frac{1}{n}\|G\|^2, \ \text{if } |\omega_x| < 1/2, \text{ otherwise } 0,$$

which agrees with the behavior of standard Monte Carlo estimate.

3.2 Analysis of Screened Poisson Reconstruction

- The MSE (variance) of the final image after screen Poisson reconstruction is given by:

$$|\epsilon_{R_\alpha}|^2 = \frac{1}{n}\frac{\alpha^4\|F\|^2 + |D(\omega_x)|^2\|G\|^2}{(\alpha^2 + |D(\omega_x)|)^2}.$$

- Using the above expression, if we set $\alpha = 0$ (i.e., using only gradient information), we have that

$$|\epsilon_{R_0}|^2 = \frac{1}{n}\frac{|D(\omega_x)|^2\|G\|^2}{(|D(\omega_x)|^2)^2} = \frac{1}{n}\frac{\|G\|^2}{|D(\omega_x)|^2} = \frac{1}{n}\frac{\|G\|^2}{2 - 2\cos(2\pi\omega_x)}.$$

This implies that the gradient construction has a singularity at $\omega_x = 0$ (the DC). However, when ω_x is high ($\omega_x = 1/2$ for example), then the variance reduces by a factor of 4. Thus, gradients are most benefitial at high frequencies.

- The optimal value of α is given by $\alpha_*^2(\omega_x) = \|G\|^2/\|F\|^2$, which gives the MSE of:

$$|\epsilon_{R_{\alpha_x}}|^2 = \frac{1}{n}\frac{\|G\|^2\|F\|^2}{\|F\|^2|D(\omega_x)|^2 + \|G\|^2}.$$

The MSE approaches $\|F\|^2/n$ at low frequency. At high frequency ($\omega_x = 1/2$), the error becomes:

$$\frac{1}{n}\frac{\|G\|^2\|F\|^2}{4\|F\|^2 + \|G\|^2} = \frac{1}{n}\frac{\|F\|^2}{4\|F\|^2/\|G\|^2 + 1} \le \frac{1}{n}\frac{\|F\|^2}{4\|F\|^2/\|G\|^2} = \frac{1}{n}\frac{\|G\|^2}{4}.$$

So, it is better than the full gradient reconstruction in this case. Thus, if $4\|G\|^2 < \|F\|^2$, then the Poisson reconstruction is better than both.

4 Gradient-Domain Path Tracing

The algorithm has already been outlined in previous sections. This section talks about the implementation details.

4.1 Symmetric Gradients

- The gradient estimate is only valid if the shift mapping is a bijection of the path space onto itself, but it is hard to find such a mapping.

- To dispense with the above requirement, they use the symmetric gradient formulation discussed in [Manzi et al., 2014].

- The symmetric gradient formulation works by computing the gradient with two integrals instead of one integral:

$$\Delta_{i,j} = \left(h(x) * \int_\Omega w_{ij}(x, \bar{p}) g_{ij}(x, \bar{p}) \ \mathrm{d}\mu(\bar{p}) \right)(x_i) - \left(h(x) * \int_\Omega w_{ji}(x, \bar{p}) g_{ji}(x, \bar{p}) \ \mathrm{d}\mu(\bar{p}) \right)(x_j).$$

The weight function w_{ij} (and w_{ji}) is set such that:

  - If both the base path and the offset path can be sampled by the path tracer, then $w_{ij}(x, \bar{p}) = 0.5$. (This means that the offset path, when evaluated in the right integral, will get the weight 0.5 as well.)
  - If the offset path cannot be sampled by the path tracer (i.e., $T_{ij}(x, \bar{p}) \notin \Omega$), we set $w_{ij} = 1$. (Here, the offset path will not show up in the right integral and so has weight 0 automatically.)

4.2 MIS

- Shift mapping can introduce high variance because the $|T'_{ij}(x, \bar{p})|$ can be come large. (On the other hand, if $|T'_{ij}(x, \bar{p})|$ is small, then the gradient is not much different from the base path's value, which means that the variance is not worse than that of naive gradient sampling.)

- $|T'_{ij}(x, \bar{p})|$ becomes large if we transition from an area where the path density is low to that where the path density is hight. In this situation, however, the mapping T_{ji} is n the exact opposite situation because the determinant would be small. As a result, it is better to use the inverse direction rather than the forward direction.

- The above thinking gives two sampling strategy for sampling the base path $\bar{x}$. First is to sample $\bar{x}$ directly. Second is to sample a path $\bar{y}$ and compute $\bar{x} = T_{ji}(\bar{y})$ and uses $\bar{x}$ as the base path. We can see that the first strategy is used in the integral to the left, and the second is used in the right.

- The best way to combine these two sampling strategy is by multiple importance sampling. Instead of letting w_{ij} take discrete values as indicated in the last section, we use the balanced heuristics and set it ot a ratio of path probabilities

$$w_{ij}(\bar{x}) = \frac{p_{ij}(\bar{x})}{p_{ij}(\bar{x}) + p_{ji}(\bar{x})}$$

where p_{ij} is the probability sampling $\bar{x}$ according to the first strategy, and p_{ji} the same thing for the second.

- Let p be the probability that the path tracer samples a path. We have that, naturally $p_{ij}(x) = p(x)$. Now,

$$p_{ji}(x) = \frac{p(y)}{|T'_{ji}(\bar{y})|}.$$

We know that $\bar{y} = T_{ij}(x)$, and, by the inverse function theorem, $|T'_{ji}(\bar{y})| = 1/T'_{ij}(\bar{x})$. So,

$$p_{ji}(x) = p(T_{ij}(\bar{x}))|T'_{ij}(\bar{x})|.$$

Thus,

$$w_{ij}(\bar{x}) = \frac{p(\bar{x})}{p(\bar{x}) + p(T_{ij}(\bar{x}))|T'_{ij}(\bar{x})|}.$$

4.3 Shift Mapping

- The energy $\|G\|^2$ affects the error (i.e., variance) of the final Poisson reconstructed image. We should minimize this energy by choosing a shift mapping that does so.

- The energy is the lowest when the path contribution of the base path and the offset path (times the determinant) is as close to each other as possible.

- Path contributions are similar when the path geometry are similar. So, the offset path should reuse as many segments of the base path as possible. The offset path has different starting point, and should hit a different point in the scene. However, after that, it should then use the base path's vertices as soon as possible.

- However, one need to be careful to choose when to connect the offset path to the base path. If the hit point of the offset path has specular or near-specular BSDF, then the change will introduce large change in BSDF values, which makes the energy goes up.

- The idea is then to connect the paths if the BSDF of the connecting segment does not include any (near-)specular vertices. Otherwise, the offset path should replicate the half-vectors of the base path.

- Specular/diffuse vertex classification is determined by a threshold of surface roughness.

- The steps are as follows:

  1. Start from the camera and shift the base ray by one pixel.

  2. If the current or the next vertex of the base path is specular, we continue to trace the offset path to avoid large changes in specular BSDFs. The outgoing direction of the offset path is determined by the half vector of the corresponding base vertex.

     At this step, we should check whether the shift is invertible:

     - The shift is not inverstible if the next base and the next offset vertex have different specular/diffuse classifications.
     - Refraction that leads to total internal reflection after the shift is also non-invertible.

     If the shift is not invertible, then the offset path should be rejected (i.e., set $w_{ij}(\bar{p})$ to 1).

  3. If the current offset vertex and the current and the next base vertex are diffuse, we connect the offset path to the base path. However, the connection segment might be occlude, in which case we reject the offset path.

References

[Kettunen et al., 2015] Kettunen, M., Manzi, M., Aittala, M., Lehtinen, J., Durand, F., and Zwicker, M. (2015). Gradient-domain path tracing. *ACM Trans. Graph.*, 34(4):123:1–123:13.

[Manzi et al., 2014] Manzi, M., Rousselle, F., Kettunen, M., Lehtinen, J., and Zwicker, M. (2014). Improved sampling for gradient-domain metropolis light transport. *ACM Trans. Graph.*, 33(6):178:1–178:12.