# Monte Carlo Techniques Basic Concepts

Chapter (13)14, 15 of "Physically Based Rendering" by Pharr&Humphreys



## Reading

- Chapter 13, 14, 15 of "Physically Based Rendering" by Pharr&Humphreys
- Chapter 7 in "Principles of Digital Image Synthesis," by A. Glassner



## Reading

| | |
|---|---|
| 13: light sources | Read on your own |
| 14.1: probability | Intro, review |
| 14.2: monte carlo | Important basics |
| 14.3: sampling random variables | Basic procedures for sampling |
| 14.4: transforming distributions | |
| 14.5: 2D sampling | |
| 15.1: Russian roulette | Improve efficiency |
| 15.2: careful sample placement | Techniques to reduce variance |
| 15.3: bias | |
| 15.4: importance sampling | |
| 15.5: sampling reflection functions | Sampling graphics |
| 15.6: sampling light sources | |
| 15.7: volume scattering | |



## Randomized Algorithms

- Las Vegas:
  - Always give right answer, but use elements of randomness on the way
  - Example: randomized quicksort
- Monte Carlo:
  - stochastic / non-deterministic
  - give the right answer on average (in the limit)



## Monte Carlo

- Efficiency, relative to other algorithms, increases with number of dimensions
- For problems such as
  - integrals difficult to evaluate because of multidimensional, complex boundary conditions (i.e., no easy closed form solutions)
  - Those with large number of coupled degrees of freedom



## Monte Carlo Integration

- Pick a set of evaluation points
- Accuracy grows with $O(N^{-0.5})$, i.e. in order to do twice as good we need 4 times as many samples
- Artifacts manifest themselves as noise
- **Research** - minimize error while minimizing the number of necessary rays

## Basic Concepts

- X, Y - random variables
  - Continuous or discrete
  - Apply function f to get Y from X: Y=f(X)
- Example - dice
  - Set of events $X_i$ = {1, 2, 3, 4, 5, 6}
  - f - rolling of dice
  - Probability of event i is $p_i = 1/6$

$$\sum_{j=1}^{6} p_j = 1$$



## Basic Concepts

- Cumulative distribution function (CDF) P(x) of a random variable X:

$$P(x) = \Pr\{X \le x\} = \int_{-\infty}^{x} p(s)ds$$

- Dice example
  - P(2) = 1/3
  - P(4) = 2/3
  - P(6)=1



## Continuous Variable

- Canonical uniform random variable $\xi$
  - Takes on all values in [0,1) with equal probability
  - Easy to create in software (pseudo-random number generator)
  - Can create general random distributions by starting with $\xi$
  - for dice example, map continuous, uniformly distributed random variable, $\xi$, to discrete random variable by choosing $X_i$ if

$$\sum_{j=1}^{i-1} p_j < \xi \le \sum_{j=1}^{i} p_j$$



## Example - lighting

- Probability of sampling illumination based on power $\Phi_i$:

$$p_i = \frac{\Phi_i}{\sum_j \Phi_j}$$

- Sums to one



## Probability Distribution Function

- Relative probability of a random variable taking on a particular value
- Derivative of CDF: $p(x) = \frac{dP(x)}{dx}$
- Non-negative
- Always integrate to one $P(x \in [a,b]) = \int_a^b p(x)dx$
- Uniform random variable:

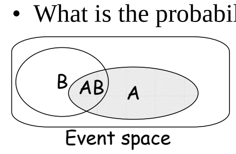
$$p(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

$$P(x) = x$$



## Cond. Probability, Independence

- We know that the outcome is in A
- What is the probability that it is in B?




Pr(B|A) = Pr(AB)/Pr(A)

- Independence: knowing A does not help: Pr(B|A) = Pr(B) ==> Pr(AB)=Pr(A)Pr(B)

## Expected Value

- Average value of the function f over some distribution of values p(x) over its domain D

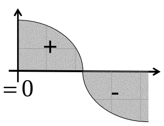
$$E_p[f(x)] = \mu = \int_D f(x)p(x)dx$$

- Example - cos over [0,π] where p is uniform

$$p(x) = 1/\pi$$

$$E_p[\cos(x)] = \int_0^{\pi} \frac{\cos x}{\pi}$$

$$= \frac{1}{\pi}(-\sin\pi + \sin 0) = 0$$






## Variance

- **Variance** of a function: expected deviation of the function from its expected value
- Fundamental concept of quantifying the error in Monte Carlo (MC) methods

$$V[f(x)] = \sigma^2 = E\left[(f(x) - \mu)^2\right]$$

- Want to reduce variance in Monte Carlo graphics algorithms



## Properties

$$E[af(x)] = aE[f(x)]$$

$$E\left[\sum_i f(X_i)\right] = \sum_i E[f(X_i)]$$

$$V[af(x)] = a^2 V[f(x)]$$

- Hence we can write:

$$V[f(x)] = E\left[(f(x))^2\right] - \mu^2$$

- For independent random variables:

$$V\left[\sum_i f(X_i)\right] = \sum_i V[f(X_i)]$$



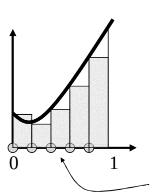
## Uniform MC Estimator

- All there is to it, really :)
- Assume we want to compute the integral of f(x) over [a,b]
- Assuming uniformly distributed random variables $X_i$ in [a,b] (i.e. p(x) = 1/(b-a))
- Our MC estimator $F_N$:

$$F_N = \frac{b-a}{N}\sum_{i=1}^{N} f(X_i)$$



## Simple Integration



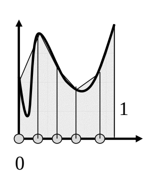

$$\int_0^1 f(x)dx \approx \sum_{i=1}^{N} f(x_i)\Delta x$$

$$= \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

$$\text{Error} = O\left(\frac{1}{N}\right)$$



## Trapezoidal Rule




$$\int_0^1 f(x)dx \approx \sum_{i=0}^{N-1}\left(f(x_i) + f(x_{i+1})\right)\frac{\Delta x}{2}$$

$$= \frac{1}{N}\sum_{i=1}^{N} w_i f(x_i)$$

$$w_i = \begin{cases} 0.5 & i = 0, N \\ 1 & 0 < i < N \end{cases}$$

$$\text{Error} = O\left(\frac{1}{N}\right)$$

## Uniform MC Estimator

- Given supply of uniform random variables: $X_i \in [a,b]$
- $E[F_N]$ is equal to the correct integral:

$$E[F_N] = E\left[\frac{b-a}{N}\sum_{i=1}^{N} f(X_i)\right] = \frac{b-a}{N}\sum_{i=1}^{N} E[f(X_i)] = \frac{b-a}{N}\sum_{i=1}^{N}\int_a^b f(x)p(x)dx = \frac{1}{N}\sum_{i=1}^{N}\int_a^b f(x)dx = \int_a^b f(x)dx$$

782 19

## General MC Estimator

- Can relax condition for general PDF
- Important for efficient evaluation of integral - draw random variable from arbitrary PDF p(X) $F_N = \frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i)}{p(X_i)}$
- And hence: $E[F_N] = E\left[\frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i)}{p(X_i)}\right] = \frac{1}{N}\sum_{i=1}^{N}\int_a^b \frac{f(x)}{p(x)}p(x)dx = \int_a^b f(x)dx$

782 20

## Confidence Interval

- We know we should expect the correct result, but how likely are we going to see it?
- Strong law of large numbers (assuming that $Y_i$ are independent and identically distributed):

$$\Pr\left\{\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} Y_i = E[Y]\right\} = 1$$

782 21

## Confidence Interval

- Rate of convergence: Chebychev Inequality

$$\Pr\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\Pr\{|F - E[F]| \geq k\} \leq \frac{V[F]}{k^2}$$

- Setting $\delta = \frac{V[F]}{k^2}$
- We have $\Pr\left\{|F - E[F]| \geq \sqrt{\frac{V[F]}{\delta}}\right\} \leq \delta$
- Answers with what probability is error below a certain amount

782 22

## MC Estimator

- How good is it? What's our error?
- Our error (root-mean square) is in the variance, hence

$$V[F_N] = V\left[\frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}\right] = \frac{1}{N^2}\sum_{i=1}^{N} V\left[\frac{f(x_i)}{p(x_i)}\right] = \frac{1}{N}V[F]$$

782 23

## MC Estimator

- Hence our overall error:

$$\Pr\left\{|F_N - E[F_N]| \geq \frac{1}{\sqrt{N}}\sqrt{\frac{V[F]}{\delta}}\right\} \leq \delta$$

- V[F] measures square of RMS error!
- This result is independent of our dimension

782 24

## Distribution of the Average

- **Central limit theorem**: sum of iid random variables with finite variance will be approximately normally distributed
- assuming normal distribution:

$$\lim_{N\to\infty} \Pr\left\{\left|F_N - E[F]\right| \le t\frac{\sigma_F}{\sqrt{N}}\right\} = \frac{1}{\sqrt{2\pi}}\int_{\infty}^{t} e^{-x^2/2}dx$$

782 25

## Distribution of the Average

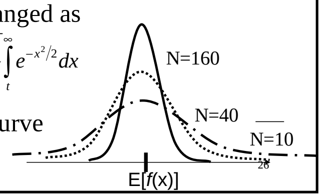
- Central limit theorem assuming normal distribution

$$\lim_{N\to\infty} \Pr\left\{\left|F_N - E[F]\right| \le t\frac{\sigma_F}{\sqrt{N}}\right\} = \frac{1}{\sqrt{2\pi}}\int_{\infty}^{t} e^{-x^2/2}dx$$

- This can be re-arranged as

$$\Pr\left\{\left|F_N - I\right| \ge t\sigma_{F_N}\right\} = \sqrt{\frac{2}{\pi}}\int_{t}^{\infty} e^{-x^2/2}dx$$

- well known Bell curve




782 26

## Distribution of the Average

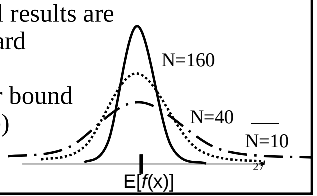
- This can be re-arranged as

$$\Pr\left\{\left|F_N - I\right| \ge t\sigma_{F_N}\right\} = \sqrt{\frac{2}{\pi}}\int_{t}^{\infty} e^{-x^2/2}dx$$

- Hence for t=3 we can conclude

$$\Pr\left\{\left|F_N - I\right| \ge 3\sigma_{F_N}\right\} = 0.997$$

- I.e. pretty much all results are within three standard deviations (probabilistic error bound - 0.997 confidence)




782 27

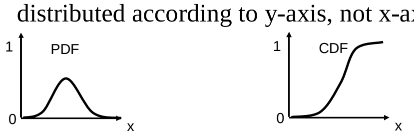
## Choosing Samples

- How to sample random variables?
- Assume we can do uniform distribution
- How to do general distributions?
  - Inversion
  - Rejection
  - Transformation

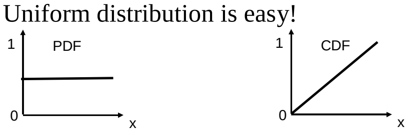
782 28

## Inversion Method

- Idea - we want all the events to be distributed according to y-axis, not x-axis




- Uniform distribution is easy!

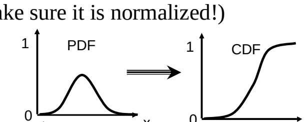



782 29

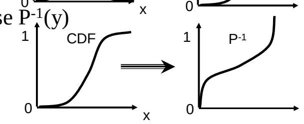
## Inversion Method

- Compute CDF (make sure it is normalized!)

$$P(x) = \int_{-\infty}^{x} p(s)ds$$




- Compute the inverse $P^{-1}(y)$




- Obtain a uniformly distributed random number $\xi$
- Compute $X_i = P^{-1}(\xi)$

782 30

## Example - Power Distribution

- Used in BSDF's $p(x) = cx^n \quad 0 \le x \le 1$
- Make sure it is normalized $\int_0^1 cx^n dx = 1 \quad c = n+1$
- Compute the CDF $P(x) = \int_0^x (n+1)s^n ds = x^{n+1}$
- Invert the CDF $P^{-1}(x) = \sqrt[n+1]{x}$
- Now we can choose a uniform ξ distribution to get a power distribution!

$$X = \sqrt[n+1]{\xi}$$

782 31

## Example - Exponential Distrib.

- E.g. Blinn's Fresnel Term $p(x) = ce^{-ax} \quad 0 \le x \le \infty$
- Make sure it is normalized $\int_0^\infty ce^{-ax} dx = 1 \quad c = a$
- Compute the CDF $P(x) = \int_0^x ae^{-as} ds = 1 - e^{-ax}$
- Invert the CDF $P^{-1}(x) = -\frac{1}{a}\ln(1-x)$
- Now we can choose a uniform x distribution to get an exponential distribution!

$$X = -\frac{1}{a}\ln(1-\xi) = -\frac{1}{a}\ln \xi$$

- extend to any funcs by piecewise approx.

782 32

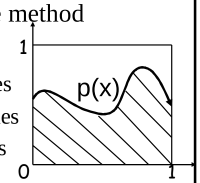
## Rejection Method

- Sometimes
  - We cannot integrate p(x)
  - We can't invert a function P(x) (we don't have the function description)
- Need to find q(x), such that p(x) < cq(x)
- Dart throwing
  - Choose a pair of random variables (X, ξ)
  - test whether ξ < p(X)/cq(X)

782 33

## Rejection Method

- Essentially we pick a point (x, ξcq(x))
- If point lies beneath p(x) then we are ok
- Not all points do -> expensive method
- Example - sampling a
  - Circle: π/4=78.5% good samples
  - Sphere: π/6=52.3% good samples
  - Gets worst in higher dimensions




782 34

## Transforming between Distrib.

- Inversion Method --> transform uniform random distribution to general distribution
- transform general X (PDF $p_x(x)$) to general Y (PDF $p_y(x)$)
- Case 1: Y=y(X)
- y(x) must be one-to-one, i.e. monotonic
- hence

$$P_y(y) = \Pr\{Y \le y(x)\} = \Pr\{X \le x\} = P_x(x)$$

782 35

## Transforming between Distrib.

- Hence we have for the PDF's:

$$p_y(y)dy = P_y(y) = P_x(x) = p_x(x)dx$$

$$p_y(y) = \left(\frac{dy}{dx}\right)^{-1} p_x(x)$$

- Example: $p_x(x) = 2x$; Y = sinX

$$p_y(y) = (\cos x)^{-1} p_x(x) = \frac{2x}{\cos x} = \frac{2\sin^{-1} y}{\sqrt{1-y^2}}$$

782 36

## Transforming between Distrib.

- y(x) usually not given
- However, if CDF's are the same, we use generalization of inversion method:

$$y(x) = P_y^{-1}(P_x(x))$$

782 37

## Multiple Dimensions

- Easily generalized - using the Jacobian of Y=T(X) $p_y(T(x)) = |J_T(x)|^{-1} p_x(x)$
- Example - polar coordinates $x = r\cos\theta$, $y = r\sin\theta$

$$J_T(x) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$$p(r,\theta) = |J_T|^{-1} p(x,y) = rp(x,y)$$

782 38

## Multiple Dimensions

- Spherical coordinates:

$$p(r,\theta,\phi) = r^2 \sin\theta p(x,y,z)$$

- Now looking at spherical directions:
- We want to solid angle to be uniformly distributed $d\omega = \sin\theta d\theta d\phi$
- Hence the density in terms of ϕ and θ:

$$p(\theta,\phi) d\theta d\phi = p(\omega) d\omega$$

$$p(\theta,\phi) = \sin\theta p(\omega)$$

782 39

## Multidimensional Sampling

- Separable case - independently sample X from $p_x$ and Y from $p_y$: $p(x,y) = p_x(x) p_y(y)$
- Often times this is not possible - compute the marginal density function p(x) first:

$$p(x) = \int p(x,y) dy$$

- Then compute conditional density function (p of y given x) $p(y \mid x) = \frac{p(x,y)}{p(x)}$
- Use 1D sampling with p(x) and p(y|x)

782 40

## Sampling of Hemisphere

- Uniformly, I.e. p(ω) = c

$$1 = \int_{H^2} p(\omega) \quad c = \frac{1}{2\pi}$$

- Sampling θ first:

$$p(\theta) = \int_0^{2\pi} p(\theta,\phi) d\phi = \int_0^{2\pi} \frac{\sin\theta}{2\pi} d\phi = \sin\theta$$

- Now sampling in ϕ:

$$p(\phi \mid \theta) = \frac{p(\theta,\phi)}{p(\theta)} = \frac{1}{2\pi}$$

782 41

## Sampling of Hemisphere

- Now we use inversion technique in order to sample the PDF's: $P(\theta) = \int_0^{\theta} \sin\alpha d\alpha = 1 - \cos\theta$

$$P(\phi \mid \theta) = \int_0^{\phi} \frac{1}{2\pi} d\alpha = \frac{\phi}{2\pi}$$

- Inverting these:

$$\theta = \cos^{-1}\xi_1$$

$$\phi = 2\pi\xi_2$$

782 42

## Sampling of Hemisphere

- Converting these to Cartesian coords:

$$\theta = \cos^{-1}\xi_1 \qquad x = \sin\theta\cos\phi = \cos(2\pi\xi_2)\sqrt{1-\xi_1^2}$$
$$\phi = 2\pi\xi_2 \qquad y = \sin\theta\sin\phi = \sin(2\pi\xi_2)\sqrt{1-\xi_1^2}$$
$$z = \cos\theta = \xi_1$$

- Similar derivation for a full sphere

782 43

## Sampling a Disk

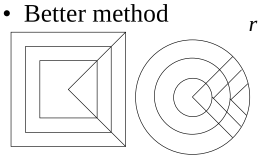
- Uniformly: $p(x,y) = \frac{1}{\pi} \quad p(r,\theta) = rp(x,y) = \frac{r}{\pi}$
- Sampling r first: $p(r) = \int_0^{2\pi} p(r,\theta)d\theta = 2r$
- Then sampling in θ: $p(\theta \mid r) = \frac{p(r,\theta)}{p(r)} = \frac{1}{2\pi}$
- Inverting the CDF: $P(r) = r^2 \quad P(\theta \mid r) = \frac{\theta}{2\pi}$

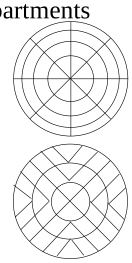
$$r = \sqrt{\xi_1} \qquad \theta = 2\pi\xi_2$$

782 44

## Sampling a Disk

- Given method distorts size of compartments
- Better method $\quad r = x \quad \theta = \frac{x}{y}$





782 45

## Cosine Weighted Hemisphere

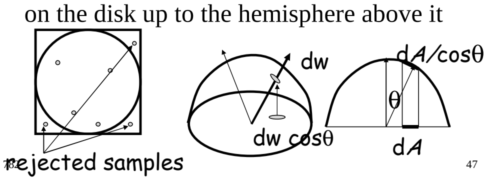
- Our scattering equations are cos-weighted!!
- Hence we would like a sampling distribution, that reflects that!
- Cos-distributed p(ω) = c.cosθ

$$1 = \int_{H^2} p(\omega)d\omega \qquad c = \frac{1}{\pi}$$
$$= \int_0^{2\pi}\int_0^{\pi/2} c\cos\theta\sin\theta d\theta d\phi \qquad p(\theta,\phi) = \frac{1}{\pi}\cos\theta\sin\theta$$
$$= 2c\pi\int_0^{\pi/2}\cos\theta\sin\theta d\theta$$

782 46

## Cosine Weighted Hemisphere

- Could use marginal and conditional densities, but use Malley's method instead:
- uniformly generate points on the unit disk
- Generate directions by projecting the points on the disk up to the hemisphere above it




782 47

## Cosine Weighted Hemisphere

- Why does this work?
- Unit disk: p(r, ϕ) = r/π
- Map to hemisphere: r = sin θ
- Jacobian of this mapping (r, ϕ) -> (sin θ, ϕ)
- Hence: $|J_T(x)| = \begin{vmatrix} \cos\theta & 0 \\ 0 & 1 \end{vmatrix} = \cos\theta$

$$p(\theta,\phi) = |J_T| p(r,\phi) = \frac{\cos\theta\sin\theta}{\pi}$$

782 48

## Performance Measure

- Key issue of graphics algorithm time-accuracy tradeoff!
- Efficiency measure of Monte-Carlo: $\varepsilon(F)=\frac{1}{V[F]T[F]}$
  - V: variance
  - T: rendering time
- Better algorithm if
  - Better variance in same time or
  - Faster for same variance
- Variance reduction techniques wanted!



## Russian Roulette

- Don't evaluate integral if the value is small (doesn't add much!)
- Example - lighting integral $L_o(p,\omega_o)=\int_{S^2} f_r(p,\omega_o,\omega_i)L_i(p,\omega_i)|\cos\theta_i|d\omega_i$
- Using N sample direction and a distribution of $p(\omega_i)$ $\frac{1}{N}\sum_{i=1}^{N}\frac{f_r(p,\omega_o,\omega_i)L_i(p,\omega_i)|\cos\theta_i|}{p(\omega_i)}$
- Avoid evaluations where $f_r$ is small or θ close to 90 degrees



## Russian Roulette

- cannot just leave these samples out
  - With some probability q we will replace with a constant c
  - With some probability (1-q) we actually do the normal evaluation, but weigh the result accordingly $F'=\begin{cases}\frac{F-qc}{1-q} & \xi>q\\ c & \text{otherwise}\end{cases}$
- The expected value works out fine $E[F']=(1-q)\left(\frac{E[F]-qc}{1-q}\right)+qc=E[F]$



## Russian Roulette

- Increases variance
- Improves speed dramatically
- Don't pick q to be high though!!



## Stratified Sampling - Revisited

- domain Λ consists of a bunch of strata $\Lambda_i$
- Take $n_i$ samples in each strata
- General MC estimator: $F_i=\frac{1}{n_i}\sum_{j=1}^{N}\frac{f(X_{i,j})}{p(X_{i,j})}$
- Expected value and variance (assuming $v_i$ is the volume of one strata): $\mu_i=E[f(X_{i,j})]=\frac{1}{v_i}\int_{\Lambda_i} f(x)dx \quad \sigma_i^2=\frac{1}{v_i}\int_{\Lambda_i}(f(x)-\mu_i)^2dx$
- Variance for one strata with $n_i$ samples: $\sigma_i^2/n_i$



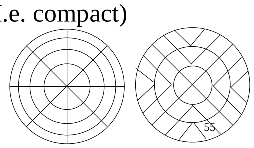
## Stratified Sampling - Revisited

- Overall estimator / variance: $V[F]=V[\sum v_iF_i]=\sum V[v_iF_i]=\sum v_i^2V[F_i]=\sum\frac{v_i^2\sigma_i^2}{n_i}$
- Assuming number of samples proportional to volume of strata - $n_i=v_iN$: $V[F_N]=\frac{1}{N}\sum v_i\sigma_i^2$
- Compared to no-strata (Q is the mean of f over the whole domain Λ): $V[F_N]=\frac{1}{N}\left(\sum v_i\sigma_i^2+\sum v_i(\mu_i-Q)\right)$

## Stratified Sampling - Revisited

$$V[F_N]=\frac{1}{N}\sum v_i\sigma_i^2 \quad V[F_N]=\frac{1}{N}\left(\sum v_i\sigma_i^2+\sum v_i(\mu_i-Q)\right)$$

- Stratified sampling never increases variance
- Right hand side minimized, when strata are close to the mean of the whole function
- I.e. pick strata so they reflect local behaviour, not global (I.e. compact)
- Which is better?

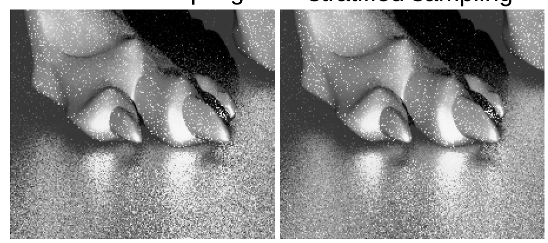




## Stratified Sampling - Revisited

- Improved glossy highlights

Random sampling stratified sampling



## Stratified Sampling - Revisited

- Curse of dimensionality
- Alternative - Latin Hypercubes
  - Better variance than uniform random
  - Worse variance than stratified



## Quasi Monte Carlo

- Doesn't use 'real' random numbers
- Replaced by low-discrepancy sequences
- Works well for many techniques including importance sampling
- Doesn't work as well for Russian Roulette and rejection sampling
- Better convergence rate than regular MC



## Bias

$$\beta = E[F]-F$$

- If β is zero - unbiased, otherwise biased
- Example - pixel filtering

$$I(x,y)=\iint f(x-s,y-t)L(s,t)\,ds\,dt$$

- Unbiased MC estimator, with distribution p

$$I(x,y)\approx\frac{1}{Np}\sum_{i=1}^{N} f(x-s_i,y-t_i)L(s_i,t_i)$$

- Biased (regular) filtering:

$$I(x,y)\approx\frac{\sum_i f(x-s_i,y-t_i)L(s_i,t_i)}{\sum_i f(x-s_i,y-t_i)}$$



## Bias

- typically $Np \neq \sum_i f(x-s_i,y-t_i)$
- I.e. the biased estimator is preferred
- Essentially trading bias for variance

## Importance Sampling MC

- Can improve our "chances" by sampling areas, that we expect have a great influence
- called "importance sampling"
- find a (known) function p, that comes close to the function we want to compute the integral of,
- then evaluate: $I = \int_0^1 p(x)\frac{f(x)}{p(x)}dx$

782 61

## Importance Sampling MC

- Crude MC: $F = \sum_{i=1}^{n} \lambda_i f(x_i)$
- For importance sampling, actually "probe" a new function f/p. I.e. we compute our new estimates to be:

$$F = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}$$

782 62

## Importance Sampling MC

- For which p does this make any sense? Well p should be close to f.
- If p = f, then we would get $F = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)} = 1$
- Hence, if we choose p = f/μ, (I.e. p is the normalized distribution function of f) then we'd get: $F = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{f(x_i)/\mu} = \mu = \int_0^1 f(x)dx$

782 63

## Optimal Probability Density

- Variance V[f(x)/p(x)] should be small
- Optimal: f(x)/p(x) is constant, variance is 0 p(x) ∝ f (x) and ∫p(x) dx = 1
- p(x) = f (x) / ∫f (x) dx
- Optimal selection is impossible since it needs the integral
- Practice: where f is large p is large

782 64

## Are These Optimal ?

$$L_r(p) = L_e(p) + \frac{R(p)}{\pi}\frac{L(p,\omega)\cos\theta}{pr(\omega)}$$

$$pr(\omega) = \frac{1}{2\pi}$$

$$pr(\omega) = \frac{\cos(\theta)}{\pi}$$

$$L_r(p) = L_e(p) + R(p)L(p,\omega)$$

782 65

## Importance Sampling MC

- Since we are finding random samples distributed by a probability given by p and we are actually evaluating in our experiments f/p, we find the variance of these experiments to be:

$$\sigma_{imp}^2 = \int_0^1 \left(\frac{f(x)}{p(x)}\right)^2 p(x)dx - I^2$$
$$= \int_0^1 \frac{f^2(x)}{p(x)}dx - I^2$$

- improves error behavior (just plug in p = f/μ)

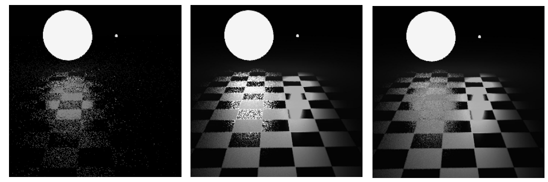
782 66

## Multiple Importance Sampling

- Importance strategy for f and g, but how to sample f*g?, e.g.
  $L_o(p,\omega_o) = \int_\Omega f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| d\omega_i$
- Should we sample according to f or according to $L_i$?
- Either one isn't good
- Use Multiple Importance Sampling (MIS)



## Multiple Importance Sampling



Importance sampling f — Importance sampling L — Multiple Importance sampling



## Multiple Importance Sampling

- In order to evaluate $\int f(x)g(x)dx$
- Pick $n_f$ samples according to $p_f$ and $n_g$ samples according to $p_g$
- Use new MC estimator:

$$\frac{1}{n_f + n_g}\left(\sum_{i=1}^{n_f} \frac{f(X_i)g(X_i)w_f(X_i)}{p_f(X_i)} + \sum_{j=1}^{n_g} \frac{f(Y_j)g(Y_j)w_g(Y_j)}{p_g(Y_j)}\right)$$

- Balance heuristic vs. power heuristic:

$$w_s(x) = \frac{n_s p_s(x)}{\sum_i n_i p_i(x)} \quad w_s(x) = \frac{(n_s p_s(x))^\beta}{\sum_i (n_i p_i(x))^\beta}$$



## MC for global illumination

- We know the basics of MC
- How to apply for global illumination?
  - How to apply to BxDF's
  - How to apply to light source



## MC for GI - general case

- General problem - evaluate:
  $L_o(p,\omega_o) = \int_\Omega f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| d\omega_i$
- We don't know much about f and L, hence use cos-weighted sampling of hemisphere in order to find a $\omega_i$
- Use Malley's method
- Make sure that $\omega_o$ and $\omega_i$ lie in same hemisphere



## MC for GI - microfacet BRDFs

- Typically based on microfacet distribution (Fresnel and Geometry terms not statistical measures)
- Example - Blinn: $D(\omega_h) = (n+2)(\omega_h \cdot N)^n$
- We know how to sample a spherical / power distribution: $\cos\theta_h = \sqrt[n+1]{\xi_1}$
  $\phi = 2\pi\xi_2$
- This sampling is over $\omega_h$, we need a distribution over $\omega_i$

## MC for GI - microfacet BRDFs

- This sampling is over $\omega_h$, we need a distribution over $\omega_i$:

$$d\omega_i = \sin\theta_i d\theta_i d\phi_i$$
$$d\omega_h = \sin\theta_h d\theta_h d\phi_h$$

- Which yields to (using that $\theta_h=2\theta_h$ and $\phi_h=\phi_h$):

$$\frac{d\omega_h}{d\omega_i} = \frac{\sin\theta_h d\theta_h d\phi_h}{\sin\theta_i d\theta_i d\phi_i} = \frac{\sin\theta_h d\theta_h d\phi_h}{\sin 2\theta_h 2 d\theta_h d\phi_h} = \frac{\sin\theta_h}{4\cos\theta_h \sin\theta_h}$$
$$= \frac{1}{4\cos\theta_h}$$

782 73

## MC for GI - microfacet BRDFs

- Isotropic microfacet model:

$$p(\theta) = \frac{p_h(\theta)}{4(\omega_o \cdot \omega_h)}$$

782 74

## MC for GI - microfacet BRDFs

- Anisotropic model (after Ashikhmin and Shirley) for a quarter disk:

$$\phi = \arctan\left(\sqrt{\frac{e_x+1}{e_y+1}}\tan\left(\frac{\pi\xi_1}{2}\right)\right)$$
$$\cos\theta_h = \xi_2^{\left(e_x\cos^2\phi + e_y\sin^2\phi + 1\right)^{-1}}$$

- If $e_x = e_y$, then we get Blinn's model

782 75

## MC for GI - Specular

- Delta-function - special treatment

$$\frac{1}{N}\sum_{i=1}^{N}\frac{f(p,\omega_i,\omega_o)L_i(p,\omega_i)|\cos\theta_i|}{p(\omega_i)} = \frac{1}{N}\sum_{i=1}^{N}\frac{\rho_{hd}(\omega_o)\frac{\delta(\omega-\omega_i)}{|\cos\theta_i|}L_i(p,\omega_i)|\cos\theta_i|}{p(\omega_i)}$$

- Since p is also a delta function

$$p(\omega_i) = \delta(\omega - \omega_i)$$

- this simplifies to

$$\rho_{hd}(\omega_o)L_i(p,\omega)$$

782 76

## MC for GI - Multiple BxDF's

- Sum up distribution densities

$$p(\omega) = \frac{1}{N}\sum_{i=1}^{N} p_i(\omega)$$

- Have three unified samples - the first one determines according to which BxDF to distribute the spherical direction

782 77

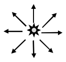
## Light Sources

- We need to evaluate
  - Sp: Cone of directions from point p to light (for evaluating the rendering equation for direct illuminations), I.e. $\omega_i$

$$L_o(p,\omega_o) = \int_\Omega f(p,\omega_i,\omega_o)L_i(p,\omega_i)|\cos\theta_i|d\omega_i$$

  - Sr: Generate random rays from the light source (Bi-directional Path Tracing or Photon Mapping)

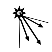
782 78



## Point Lights

- Source is a point
- uniform power in all directions
- hard shadows
- Sp:
  - Delta-light source
  - Treat similar to specular BxDF
- Sr: sampling of a uniform sphere





## Spot Lights

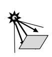
- Like point light, but only emits light in a cone-like direction
- Sp: like point light, i.e. delta function
- Sr: sampling of a cone

$$p(\theta,\phi) = p(\theta)p(\phi) \qquad 1 = c\int_0^{\theta_{\max}} \sin\theta d\theta = c(1-\cos\theta_{\max})$$

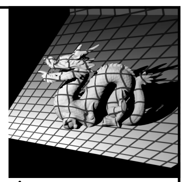
$$p(\phi) = 1/2\pi \qquad p(\theta) = 1/(1-\cos\theta_{\max})$$

$$p(\theta) = c$$





## Projection Lights



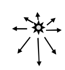
- Like spot light, but with a texture in front of it
- Sp: like spot light, i.e. delta function
- Sr: like spot light, i.e. sampling of a cone

$$p(\theta,\phi) = p(\theta)p(\phi) \qquad 1 = c\int_0^{\theta_{\max}} \sin\theta d\theta = c(1-\cos\theta_{\max})$$

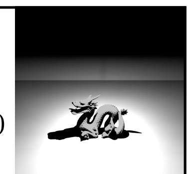
$$p(\phi) = 1/2\pi \qquad p(\theta) = 1/(1-\cos\theta_{\max})$$

$$p(\theta) = c$$

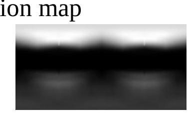




## Goniophotometric Lights

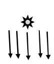


- Like point light (hard shadows)
- Non-uniform power in all directions - given by distribution map

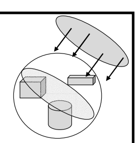


- Sp: like point-light
  - Delta-light source
  - Treat similar to specular BxDF
- Sr: like point light, i.e. sampling of a uniform sphere

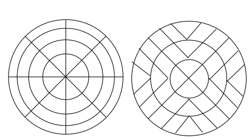




## Directional Lights

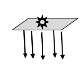


- Infinite light source, i.e. only one distinct light direction
- hard shadows
- Sp: like point-light
  - Delta function



- Sr:
  - create virtual disk of the size of the scene
  - sample disk uniformly (e.g. Shirley)





## Area Lights

- Defined by shape
- Soft shadows
- Sp: distribution over solid angle
  - $\theta_o$ is the angle between $\omega_i$ and (light) shape normal
  - A is the area of the shape

$$d\omega_i = \frac{\cos\theta_o}{r^2} dA$$

- Sr:
  - Sampling over area of the shape
- Sampling distribution depends on the area of the shape

$$p(x) = \frac{1}{A}$$

## Area Lights

- If v(p,p') determines visibility:

$$L_o(p,\omega_o) = \int_\Omega f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| d\omega_i$$

$$= \int_A v(p,p') f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| \frac{\cos\theta_o}{r^2} dA$$

$$d\omega_i = \frac{\cos\theta_o}{r^2} dA$$

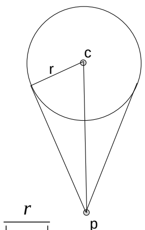
- Hence: $p(x) = \frac{1}{A}$

$$L_o(p,\omega_o) \approx \frac{1}{p(x)} v(p,p') f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| \frac{\cos\theta_o}{r^2}$$

$$\approx \frac{A}{r^2} v(p,p') f(p,\omega_i,\omega_o) L_i(p,\omega_i) |\cos\theta_i| \cos\theta_o$$

782 85

## Spherical Lights




- Special area shape
- Not all of the sphere is visible from outside of the sphere
- Only sample the area, that is visible from p
- Sp: distribution over solid angle
  - Use cone sampling $\sin\theta_{max} = \frac{r}{|p-c|}$
- Sr: Simply sample a uniform sphere

782 86

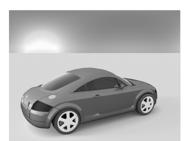
## Infinite Area Lights

- Typically environment light (spherical)
- Encloses the whole scene
- Sp:
  - Normal given - cos-weighted sampling
  - Otherwise - uniform spherical distribution
- Sr:
  - uniformly sample sphere at two points $p_1$ and $p_2$
  - The direction $p_1$-$p_2$ is the uniformly distributed ray

782 87

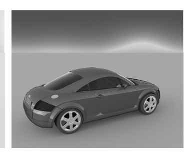
## Infinite Area Lights

Area light + directional light



Morning skylight

Midday skylight



Sunset environment map

782 88

## Summary

$$E_p[f(x)] = \mu = \int_D f(x) p(x) dx$$

$$V[f(x)] = \sigma^2 = E\left[(f(x)-\mu)^2\right]$$

$$F_N = \frac{1}{N}\sum_{i=1}^{N} \frac{f(X_i)}{p(X_i)}$$

$$F_N = \frac{b-a}{N}\sum_{i=1}^{N} f(X_i)$$

$$V[F_N] = \frac{1}{N} V[F]$$

782 89