Overview

Earlier lecture
- Statistical sampling and Monte Carlo integration

Last lecture
- Signal processing view of sampling

Today
- Variance reduction
- Importance sampling
- Stratified sampling
- Multidimensional sampling patterns
- Discrepancy and Quasi-Monte Carlo

Latter
- Path tracing for interreflection
- Density estimation

Cameras

\[ R = \int \int \int \int L(x, \omega, t) \ P(x) \ S(t) \ \cos \theta \ dA \ d\omega \ dt \]

Motion Blur

Depth of Field

Source: Cook, Porter, Carpenter, 1984
Source: Mitchell, 1991
### Variance

**Definition**

\[ V[Y] \equiv E[(Y - E[Y])^2] \]
\[ = E[Y^2] - E[Y]^2 \]

**Variance decreases with sample size**

\[ V[\frac{1}{N} \sum_{i=1}^{N} Y_i] = \frac{1}{N^2} \sum_{i=1}^{N} V[Y_i] = \frac{1}{N^2} NV[Y] = \frac{1}{N} V[Y] \]
**Variance Reduction**

Efficiency measure

\[
\text{Efficiency} \propto \frac{1}{\text{Variance} \cdot \text{Cost}}
\]

If one technique has twice the variance as another technique, then it takes twice as many samples to achieve the same variance.

If one technique has twice the cost of another technique with the same variance, then it takes twice as much time to achieve the same variance.

Techniques to increase efficiency
- Importance sampling
- Stratified sampling

---

**Biasing**

Previously used a uniform probability distribution.

Can use another probability distribution

\[
X_i \sim p(x)
\]

But must change the estimator

\[
Y_i = \frac{f(X_i)}{p(X_i)}
\]
**Unbiased Estimate**

**Probability** \( X_i \sim p(x) \)

**Estimator** \( Y_i = \frac{f(X_i)}{p(X_i)} \)

\[
E[Y_i] = E \left[ \frac{f(X_i)}{p(X_i)} \right] \\
= \int \frac{f(X_i)}{p(X_i)} p(x) \, dx \\
= \int f(x) \, dx \\
= 1
\]

**Importance Sampling**

**Sample according to** \( f \)

\[
\tilde{p}(x) = \frac{f(x)}{E[f]} \\
\int \tilde{p}(x) \, dx = \int \frac{f(x)}{E[f]} \, dx \\
= \frac{1}{E[f]} \int f(x) \, dx \\
= 1
\]
Importance Sampling

Variance

\[ V[f] = E[f^2] - E^2[f] \]

Sample according to \( f \)

\[ \tilde{p}(x) = \frac{f(x)}{E[f]} \]

\[ \tilde{f}(x) = \frac{f(x)}{\tilde{p}(x)} \]

Zero variance!

\[ V[\tilde{f}^2] = 0 \]

Examples

Projected solid angle

- 4 eye rays per pixel
- 100 shadow rays

Area

- 4 eye rays per pixel
- 100 shadow rays
Irradiance

Generate cosine weighted distribution

\[ p(\omega) d\omega = \cos \theta d\omega \]

\[ E = \int_{\mathbb{H}^2} L_i(\omega_i) \cos \theta_i d\omega_i \]

Stratified Sampling

Stratified sampling is like jittered sampling

Allocate samples per region

\[ F_N = \frac{1}{N} \sum_{i=1}^{N} F_i \]

New variance

\[ V[F_N] = \frac{1}{N^2} \sum_{i=1}^{N} V[F_i] \]

Thus, if the variance in regions is less than the overall variance, there will be a reduction in resulting variance

For example: An edge through a pixel

\[ V[F_N] = \frac{1}{N^2} \sum_{i=1}^{N} V[F_i] = \frac{V[F_E]}{N^{1.5}} \]
Sampling a Circle

Equi-Areal

\[ \theta = 2\pi U_1 \]
\[ r = \sqrt{U_2} \]

Shirley’s Mapping

\[ r = U_1 \]
\[ \theta = \frac{\pi U_2}{4 U_1} \]
High-dimensional Sampling

Numerical quadrature
For a given error ...

\[ E \sim 1/n = \frac{1}{N^{1/d}} \]

Random sampling
For a given variance ...

\[ E \sim V^{1/2} \sim \frac{1}{N^{1/2}} \]

---

Monte Carlo requires fewer samples for the same error in high dimensional spaces

Block Design

Latin Square

<table>
<thead>
<tr>
<th>a</th>
<th>d</th>
<th>c</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>a</td>
<td>d</td>
<td>c</td>
</tr>
<tr>
<td>c</td>
<td>b</td>
<td>a</td>
<td>d</td>
</tr>
<tr>
<td>d</td>
<td>c</td>
<td>b</td>
<td>a</td>
</tr>
</tbody>
</table>

Alphabet of size \( n \)
Each symbol appears exactly once in each row and column
Rows and columns are stratified
Block Design

N-Rook Pattern

Incomplete block design

Replaced $n^2$ samples with $n$ samples

Permutations: $(\pi_1(i), \pi_2(i), \ldots, \pi_d(i))$

Generalizations: N-queens, 2D projection

$(\pi_x = \{1, 2, 3, 4\}, \pi_y = \{4, 2, 3, 1\})$

Space-time Patterns

Distribute samples in time

- Complete in space
- Samples in space should have blue-noise spectrum
- Incomplete in time
- Decorrelate space and time
- Nearby samples in space should differ greatly in time

Cook Pattern

<table>
<thead>
<tr>
<th>6</th>
<th>10</th>
<th>2</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>14</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Pan-diagonal Magic Square

<table>
<thead>
<tr>
<th>15</th>
<th>8</th>
<th>5</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Path Tracing

4 eye rays per pixel
16 shadow rays per eye ray

64 eye rays per pixel
1 shadow ray per eye ray

Complete

Incomplete
Discrepancy

\[ \Delta(x, y) = \frac{n(x, y)}{N} - xy \]

\[ A = xy \]

\[ n(x, y) \] number of samples in \( A \)

\[ D_N = \max_{x,y} |\Delta(x, y)| \]

Theorem on Total Variation

**Theorem:**

\[ \left| \frac{1}{N} \sum_{i=1}^{N} f(X_i) - \int f(x) \, dx \right| \leq V(f)D_N \]

**Proof: Integrate by parts**

\[ \int f(x) \left( \frac{\delta(x - x_i)}{N} - 1 \right) \, dx = \int f(x) \frac{\partial \Delta(x)}{\partial x} \, dx \]

\[ = \frac{\partial \Delta(x)}{\partial x} \Delta(x) \, dx = -\int \frac{\partial f(x)}{\partial x} \Delta(x) \, dx \]

\[ \leq D_N \int \left| \frac{\partial f(x)}{\partial x} \right| \, dx = V(f)D_N \]
Quasi-Monte Carlo Patterns

Radical inverse (digit reverse)
of integer $i$ in integer base $b$

$i = d_i \cdots d_2 d_1 d_0$
$\phi_b(i) \equiv 0.d_0 d_1 d_2 \cdots d_i$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\phi_2(i)$</th>
<th>$\phi_3(i)$</th>
<th>$\phi_5(i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>.1</td>
<td>1/2</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>.01</td>
<td>1/4</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>.11</td>
<td>3/4</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>.001</td>
<td>3/8</td>
</tr>
<tr>
<td>5</td>
<td>101</td>
<td>.101</td>
<td>5/8</td>
</tr>
</tbody>
</table>

Hammersley points

$(i / N, \phi_2(i), \phi_3(i), \phi_5(i), \cdots)$

$D_N = O\left(\frac{\log^{d-1} N}{N}\right)$

Halton points (sequential)

$(\phi_2(i), \phi_3(i), \phi_5(i), \cdots)$

$D_N = O\left(\frac{\log^d N}{N}\right)$

Hammersley Points

$(i / N, \phi_2(i), \phi_3(i), \phi_5(i), \cdots)$
### Edge Discrepancy

\[ ax + by + c \]

Note: SGI IR Multisampling extension:
8x8 subpixel grid; 1,2,4,8 samples

### Low-Discrepancy Patterns

<table>
<thead>
<tr>
<th>Process</th>
<th>16 points</th>
<th>256 points</th>
<th>1600 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zaremba</td>
<td>0.0504</td>
<td>0.00478</td>
<td>0.00111</td>
</tr>
<tr>
<td>Jittered</td>
<td>0.0538</td>
<td>0.00595</td>
<td>0.00146</td>
</tr>
<tr>
<td>Poisson-Disk</td>
<td>0.0613</td>
<td>0.00767</td>
<td>0.00241</td>
</tr>
<tr>
<td>N-Rooks</td>
<td>0.0637</td>
<td>0.0123</td>
<td>0.00488</td>
</tr>
<tr>
<td>Random</td>
<td>0.0924</td>
<td>0.0224</td>
<td>0.00866</td>
</tr>
</tbody>
</table>

Discrepancy of random edges, From Mitchell (1992)

- Random sampling converges as \( N^{-1/2} \)
- Zaremba converges faster and has lower discrepancy
- Zaremba has a relatively poor blue noise spectra
- Jittered and Poisson-Disk recommended
Views of Integration

1. Signal processing
   - Sampling and reconstruction, aliasing and antialiasing
   - Blue noise good
2. Statistical sampling (Monte Carlo)
   - Sampling like polling
   - Variance
   - High dimensional sampling: $1/N^{1/2}$
3. Quasi Monte Carlo
   - Discrepancy
   - Asymptotic efficiency in high dimensions
4. Numerical
   - Quadrature/Integration rules
   - Smooth functions