

# Monte Carlo integrování

Martin Bulant

7. listopadu 2011

## Opakování

Při renderingu se řeší zobrazovací rovnice, kterou jsme si odvodili na minulé přednášce. Zobrazovací rovnice popisuje ustálený stav, neboli energetickou rovnováhu světla ve scéně. Tato rovnováha platí pro všechny reálné scény. Zobrazovací rovnice je komplexní matematická formulace problému.

## Tvary zobrazovací rovnice

### Plošný tvar

$$L_o(x, \omega_o) = L_e(x, \omega_o) + \int_M L_o(y \rightarrow x) \cdot f_r(y \rightarrow x \rightarrow \omega_o) \cdot G(x \leftrightarrow y) \cdot V(x \leftrightarrow y) dA_y$$

### Úhlový tvar

$$L_o(x, \omega_o) = L_e(x, \omega_o) + \int_H L_o(r(x, \omega_i), -\omega_i) \cdot f_r(x, \omega_i \rightarrow \omega_o) \cdot \cos(\theta_i) d\omega_i,$$

kde  $r(x, \omega_i)$  je průsečík paprsku z bodu  $x$  ve směru  $\omega_i$  se scénou.

## Řešení zobrazovací rovnice

Zobrazovací rovnici lze řešit mnoha způsoby. Například radiační metodou, která ovšem využívá určité předpoklady, jako například, že všechny plochy jsou Lambertovské (mají konstantní brdf). Formálně je zobrazovací rovnice integrální rovnicí, kterou neumíme analyticky vyřešit.

## Operátorový tvar zobrazovací rovnice

Zobrazovací rovnici lze přepsat do tvaru:  $L = L_e + T \circ L$ , Kde  $T$  je operátor přenosu a  $L$  je rozložení radiance ve scéně v rovnovážném stavu. Jedna aplikace operátoru  $T$  z jednoho rozložení radiance ve scéně udělá jiné rozložení odpovídající jednomu odrazu světla. Pokud už se aplikováním operátoru  $T$  rozložení radiance ve scéně nemění, dostali jsme se do pevného bodu, tedy k řešení zobrazovací rovnice. Tento zápis velmi zjednodušuje celou rovnici a umožňuje

tak lépe nahlédnout možné způsoby řešení. Prvním je rekurzivní řešení. Druhým je integrování přes nezávislé cesty.

## Monte Carlo integrování

Jelikož při řešení zobrazovací rovnice se řeší integrály, které jsou nespojité a mají téměř libovolné hodnoty, nelze úplně jednoše použít kvadraturní vzorce pro numerický výpočet hodnoty integrálu. Z tohoto důvodu se používá obecnější postup, kterým je Monte Carlo integrování.

Kvadraturní vzorce navíc nejsou užitečné pro vyšší dimenze, jelikož rychlost jejich konvergence závisí na dimenzi prostoru ve kterém se integruje.

Na rozdíl od kvadraturních vzorců se v MC metodách vzorky rozmísťují náhodně popřípadě pseudonáhodně. Rychlost konvergence MC závisí pouze na počtu vzorků a je  $O(\frac{1}{\sqrt{N}})$ . Je tedy rychlejší než kvadraturní vzorce pro 3 a více dimenzí. Při použití speciálních metod pro rozmísťování vzorků lze rychlost konvergence ještě zvýšit. Jde například o metodu quazi-Monte Carlo.

## Náhodné veličiny

### Diskrétní

Diskrétní náhodné veličiny mají konečný počet různých hodnot, které nabývají s určitou pravděpodobností. Rozložení pravděpodobnosti určuje, s jakou pravděpodobností jsou nabývány které hodnoty. To lze zapsat pomocí pravděpodobnostní funkce. Toto rozložení lze odvodit i z distribuční funkce, která pro danou hodnotu k její pravděpodobnosti přičte součet pravděpodobností všech předchozích hodnot. Distribuční funkce je neklesající.

### Spojité

Spojité náhodné veličiny jsou zobecněním diskretních náhodných veličin pro nespočetné prostory. Pro spojité náhodné veličiny se místo pravděpodobnostní funkce používá hustota pravděpodobnosti  $p(x)$  (probability density function, **pdf**). Pravděpodobnost  $\Pr(X \in D)$  se pomocí ní spočítá takto:

$$\Pr(X \in D) = \int_D p(x) dx$$

Název distribuční funkce  $P(x)$  (cumulative distribution function, **cdf**) je zachován, jen je v definici místo součtu použit integrál:

$$P(x) \equiv \Pr(X \leq x) = \int_{-\infty}^x p(t) dt$$

### Rovnoměrné rozložení

Pro rovnoměrné rozdělení je  $pdf = \frac{1}{S}$ , kde  $S$  je velikost té oblasti přes kterou integruji. Pro hemisféru je  $pdf = \frac{1}{\pi}$

### Vlastnosti náhodné veličiny

Náhodné veličiny jsou charakterizovány agregátními veličinami jako je střední hodnota a rozptyl. Střední hodnota  $E[X]$  je integrál hodnoty náhodné veličiny vážený hustotou pravděpodobnosti:

$$E[X] = \int_D xp(x) dx$$

Rozptyl  $V[X]$  se počítá dle vzorce:

$$V[X] = E[X^2] - E[X]^2$$

Pro nezávislé náhodné veličiny  $X_i$  platí:

$$V\left[\sum_i X_i\right] = \sum_i V[X_i]$$

Dále platí, že rozptyl není lineární, tedy:

$$V[aX] = a^2V[X]$$

Pokud  $Y$  vznikne transformací náhodné veličiny  $X$  funkcí  $f$ , pak  $Y$  je též náhodná veličina a pro její střední hodnotu platí vzorec:

$$E[Y] = \int_D f(x)p(x) dx$$

## Odhad integrálu metodou Monte Carlo

Budeme odhadovat integrál  $I$  nějaké funkce  $f(x)$  definovaný takto:

$$I = \int_{\Omega} f(x) dx$$

### Primární estimátor

Estimátor je náhodná veličina, která vznikla určitou transformací jiné náhodné veličiny. Její realizace (hodnota) je konkrétní odhad. Primární estimátor  $Y$  integrálu  $I$  definujeme vzorcem

$$F_{\text{prim}} = Y = \frac{f(X)}{p(X)},$$

kde  $X$  je náhodná veličina s hustotou pravděpodobnosti  $p(x)$ . Toto funguje pro libovolnou hustotu pravděpodobnosti. Pro jiné než rovnoměrné hustoty pravděpodobnosti ale potřebujeme funkci, která umí generovat náhodné vzorky s daným rozdělením a navíc pro daný vzorek musíme být schopni vyhodnotit hustotu pravděpodobnosti, s jakou byl vygenerován.

### Nestrannost primárního estimátoru

Estimátor je nestranný pokud v průměru dává správnou veličinu, tedy je bez systematické chyby, to je vyjádřeno vzorcem:

$$E[F] = Q,$$

kde  $F$  je estimátor náhodné veličiny  $Q$  a  $Q$  je odhadovaná veličina.

Nestrannost je velmi výhodná vlastnost, jelikož jediný zdroj chyby je náhodná oscilace okolo hodnoty  $Q$ . Nestrannost primárního estimátoru plyne z rovnic:

$$E[F_{\text{prim}}] = E[Y] = \int_{\Omega} \frac{f(x)}{p(x)} p(x) dx = I,$$

kde  $\Omega$  je oblast kterou chci integrovat. Platí důležitá podmínka:

$$\forall x \in X : f(x) \neq 0 \implies p(x) \neq 0$$

### Rozptyl primárního estimátoru

Rozptyl neboli variance je jediné měřítko pro určení chyby nestranného estimátoru.

$$V[F_{\text{prim}}] = \sigma_{\text{prim}}^2 = \int_{\Omega} \frac{f(x)^2}{p(x)} dx - I^2$$

Pokud  $p(x)$  bude podobná  $f(x)$  bude výsledný rozptyl malý.

### Sekundární estimátor

Jelikož použití pouze jednoho vzorku v primárním estimátoru nezaručuje dostatečně malý rozptyl, používá se estimátor sekundární. Ten využívá  $N$  nezávislých náhodných veličin  $Y_i = f(X_i)/p(X_i)$  a jako výsledek se vezme jejich průměr, tedy:

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

Tento estimátor je opět nestranný, jelikož platí:

$$E[F_N] = E\left[\frac{1}{N} \sum Y_i\right] = \frac{1}{N} \sum E[Y_N] = \frac{1}{N} NI = I$$

Hustota pravděpodobnosti obecně nemusí být identická pro všechny vzorky.

## Rozptyl sekundárního estimátoru

Pro rozptyl platí vzorec:

$$V[F_N] = \frac{1}{N} V[F_{\text{prim}}]$$

Rozptyl je tedy přímo úměrný rozptylu primárního estimátoru s koeficientem  $\frac{1}{N}$ . Standardní odchylka je definována jako odmocnina z rozptylu a proto je pouze  $\sqrt{N}$ -krát menší.

## Vzorkování podle důležitosti (Importance sampling)

Části vzorkovaného intervalu, kde má  $f$  větší hodnotu, jsou důležitější, protože vzorky z těchto oblastí více ovlivňují výsledek. Vzorkování podle důležitosti umisťuje vzorky přednostně do takových oblastí. Toho se docílí tím, že pdf ze které se vybírají vzorky bude podobná integrandu. Použitím vzorkování důležitosti lze snížit rozptyl při zachování nestrannosti.

Pokud bychom použili pdf přímo úměrnou  $f(x)$ , tedy bychom měli

$$p(x) = \frac{f(x)}{N},$$

kde  $N$  je normalizační faktor definovaný jako

$$N = \int_{\Omega} f(x) dx,$$

který zajišťuje, že se  $p(x)$  integruje do 1 a tedy že  $p(x)$  je hustota pravděpodobnosti, pak by rozptyl byl nulový jak je vidět dále.

$$V[F_{\text{prim}}] = \int_{\Omega} \frac{f(x)^2}{p(x)} dx - I^2 = \int_{\Omega} f(x) \cdot N dx - I^2 = N \int_{\Omega} f(x) dx - I^2$$

Po dosazení za  $N$  a použití definice námi odhadovaného integrálu  $I = \int_{\Omega} f(x) dx$  dostaneme:

$$= \left( \int_{\Omega} f(x) dx \right)^2 - I^2 = I^2 - I^2 = 0$$

Bohužel v praxi se taková pdf nedá použít, protože konstrukce  $p(x)$  přímo úměrné  $f(x)$  vyžaduje normalizaci, tedy integrál, který se snažíme spočítat.

## Odhad irradiance

Pro výpočet irradiance v bodě  $x$  platí následující rovnice:

$$E(x) = \int_{H(x)} L_i(x, \omega_i) \cdot \cos \theta_i d\omega_i,$$

kde  $H(x)$  je hemisféra nad bodem  $x$ .

## Uniformní vzorkování

Při uniformním vzorkování hemisféry máme  $p(\omega) = \frac{1}{2\pi}$ . Výsledný estimátor má pak tvar

$$F_N = \frac{2\pi}{N} \sum_{k=1}^N L_i(x, \omega_{i,k}) \cdot \cos \theta_{i,k},$$

kde  $\omega_{i,k}$  je k-tý vygenerovaný směr,  $\theta_{i,k}$  je úhel, který svírá normála v bodě  $x$  a směr  $\omega_{i,k}$ .

## Vzorkování dle cosinu

Pro vzorkování dle funkce cosinus je třeba použít  $p(\omega) = \frac{\cos \theta}{\pi}$ . Výsledný estimátor má pak tvar

$$F_N = \frac{\pi}{N} \sum_{k=1}^N L_i(x, \omega_{i,k}),$$

kde  $\omega_{i,k}$  je k-tý vygenerovaný směr.

## Výpočet osvětlení z jednoho zdroje světla

Při výpočtu osvětlení z jednoho zdroje světla můžeme použít obě předchozí metody. V obou případech je  $L_i$  buď nula nebo rovna emitované radianci. V případě využití uniformního vzorkování se generuje hodně vzorků a násobí se malým číslem tam, kde je hodnota  $\cos$  malá. Budeme tedy vzorky generovat pomocí hustoty pravděpodobnosti úměrné funkci  $\cos$ , tím použijeme vzorkování dle cosinu. Obě metody budou v průměru dávat stejný výsledek. Výsledek bude obsahovat mnoho šumu, jelikož pravděpodobnost trefení světla je velmi malá. Pro některé pixely se trefíme, pro jiné ne. Obě metody mají ale stejnou nevýhodu, a to, že generují mnoho vzorků, které se poté násobí hodnotou 0, protože se netrefí světlo. To se ovšem neděje při vzorkování zdroje světla

## Vzorkování zdroje světla

Při vzorkování zdroje světla využijeme pro výpočet irradiance plošný tvar zobrazovací rovnice. Budeme tedy řešit rovnici:

$$E(x) = \int_A L_e(y \rightarrow x) \cdot G(y \leftrightarrow x) \cdot V(y \leftrightarrow x) dA,$$

kde  $y$  je bod na zdroji světla a střed  $dA$ .

Budeme integrovat MC a vzorkovat uniformně plochu zdroje,  $p(y)$  tedy bude rovno  $\frac{1}{|A|}$  a výsledný estimátor bude mít tvar:

$$F_N = \frac{|A|}{N} \sum_{k=1}^N L_e(y_k \rightarrow x) \cdot G(y_k \leftrightarrow x) \cdot V(y_k \leftrightarrow x) dA$$

Náhodně tedy vybereme  $y$  na zdroji světla, z popisu scény získáme radianci ve směru  $x$ , otestujeme viditelnost a přenásobíme geometrickým faktorem. V průměru opět dostaneme stejný odhad jako v předchozích 2 případech. Ale množství šumu bude menší. Do tohoto přístupu nejde jednoduše zabudovat cosinový faktor. Naštěstí to tolik nevádí, jelikož velikost průmětu zdroje světla na jednotkovou polokouli je malá a rozdíl hodnot funkce  $\cos$  pro různé vzorky na tomto zdroji je tedy zanedbatelný.

### Přímé osvětlení na ploše s různou BRDF

V tomto případě nás nezajímá irradiance, ale odražená radiance. Opět se použije vzorkování zdroje světla. Odhadujeme integrál:

$$L_o(x, \omega_o) = \int_M L_e(y \rightarrow x) \cdot f_r(y \rightarrow x \rightarrow \omega_o) \cdot G(y \leftrightarrow x) \cdot V(y \leftrightarrow x) dA_y$$

Pro odhad tohoto integrálu se použije estimátor

$$F_N = \frac{|A|}{N} \sum_{k=1}^N L_e(y_k \rightarrow x) \cdot f_r(y_k \rightarrow x \rightarrow \omega_o) \cdot G(y_k \leftrightarrow x) \cdot V(y_k \leftrightarrow x)$$

### Nepřímé osvětlení na ploše s obecnou BRDF

Chceme získat odchozí světlo, což je integrál příchozího světla ze všech směrů. Odhadujeme integrál:

$$L_o^{ind}(x, \omega_o) = \int_{H(x)} L_r(r(x, \omega_i), -\omega_i) \cdot f_r(x, \omega_i \rightarrow \omega_o) \cdot \cos(\theta_i) d\omega_i,$$

kde  $L_r$  je odražené světlo ze vzorce  $L_o = L_r + L_e$  v bodě  $r(x, \omega_i)$ ,  $r(x, \omega_i)$  je bod, kde paprsek z bodu  $x$  ve směru  $\omega_i$  protne scénu. Odpovídající estimátor tedy je:

$$L_o^{ind}(x, \omega_o) = \frac{|A|}{N} \sum_{k=1}^N \frac{L_r(r(x, \omega_{i,k}), -\omega_{i,k}) \cdot f_r(x, \omega_{i,k} \rightarrow \omega_o) \cdot \cos(\theta_{i,k})}{p(\omega_{i,k})},$$

## Výpočet osvětlení ve scéně

### Distribution raytracing

Přímé osvětlení počítá vzorkováním zdrojů světla. Pro nepřímé střílí více sekundárních paprsků. To způsobuje velkou explozi počtu vzorků s hloubkou rekurze. Problémem je, že nejvíc vzorků se investuje tam, kde je velmi malá energie a tedy malé příspěvky. Další nevýhodou je ad-hoc ukončení rekurze dosažením maximální hloubky nebo minimálního příspěvku.

## Sledování cest (Path tracing)

Narozdíl od Distribution raytracingu používá pouze 1 sekundární paprsek, kdy se po dopadu paprsku do scény vebere způsob interakce v daném místě a pak se provádí vzorkování podle důležitosti té vybrané interakce. Pro výpočet přímého osvětlení lze čekat než náhodně vygenerovaný paprsek trefí zdroj nebo použít nějakou metodu pro výpočet přímého osvětlení v každém vrcholu procházené cesty. Výsledná hodnota pixelu bude průměr z mnoha cest z kamery přes daný pixel.

### Sledování cest jako řešení zobrazovací rovnice metodou MC

Na sledování cest se lze dívat jako na sekundární estimátor, kde náhodnou veličinou je cesta ve scéně a integrovanou funkcí je hodnota příchozí radiance z dané cesty přenásobená hodnotou BRDF a cosinovým faktorem.

$$L(x, \omega) = \int_{c \in M} L(c) \cdot f_r(c, x, \omega) \cdot \cos(\theta) dc,$$

kde  $M$  je prostor cest všech délek ve scéně,  $f_r(c, x, \omega)$  je hodnota BRDF v bodě  $x$  pro směr  $\omega$  a směr, kterým pokračuje cesta  $c$ , a  $\theta$  je úhel, který svírá normála se směrem, kterým pokračuje cesta  $c$ . Integruje se tedy přes spoustu dimenzí v prostoru cest.

### Ukončení rekurze ruskou ruletou

Sledování cest umožňuje ukončit rekurzi korektně za použití tzv. ruské rulety. Kdy se v rekurzi pokračuje s pravděpodobností  $q$  a váha vzorku, který pokračuje, se upraví koeficientem  $1/q$ . Formálně jde o transformace náhodné veličiny dle vzorce

$$Z = \begin{cases} Y/q & \text{pokud } \xi \leq q \\ 0 & \text{jinak} \end{cases},$$

kde  $\xi$  je náhodné číslo z intervalu  $[0, 1]$  a  $q$  pravděpodobnost ukončení rekurze. Přičemž takováto transformace zachovává nestrannost, jak je vidět zde:

$$E[Z] = \frac{E[Y]}{q} \cdot q + 0 \cdot \frac{1}{q-1} = E[Y]$$

## Reference

- [1] Krivánek, J.: Monte Carlo integrování. 2011.  
URL <http://cgg.mff.cuni.cz/~jaroslav/teaching/2011-pg3/slides/krivanek-06-npgr010-2011-mc1.pdf>