

# Počítačová grafika III (NPGR010)

## Přednáška: Monte Carlo integrování, estimátory

Hubert Kindermann

### 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.

### 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.

### 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. Numerické hodnoty jednotlivých integrálů v naší rovnici nelze dobře odhadovat pomocí kvadratických vzorců, z důvodu nespojitosti a vysokého rozsahu integrovancých funkcí. Kvadratické vzorce navíc nejsou užitečné pro vyšší dimenze, jelikož rychlost jejich konvergence závisí na dimenzi prostoru ve kterém se integruje. Rychlost jejich konvergence lze vyjádřit jako

$$O(N^{-\frac{1}{s}}),$$

kde  $s$  je dimenze prostoru a  $N$  je počet vzorkovacích bodů. Proto tyto odhady budeme hledat metodou Monte Carlo, jejíž rychlost konvergence na dimenzi problému nezáleží a kterou v následujícím textu popíšeme.

### Monte Carlo integrování

Metody typu Monte Carlo jsou takové metody, které využívají opakovaného vzorkování funkce v náhodných bodech, aby určily výsledek. Při použití pro odhad hodnoty integrálu funkce, vyhodnocujeme integrovanou funkci v náhodných bodech a snažíme se z těchto vzorků odhadnout hodnotu výsledku. Rychlost konvergence MC integrace závisí pouze na počtu vzorků a je

$$O(N^{-\frac{1}{2}}).$$

Je tedy rychlejší než kvadratické vzorce pro 3 a více dimenzí. Při použití speciálních metod pro rozmístění vzorků lze rychlost konvergence ještě zvýšit. Jde například o metody quazi-Monte Carlo.

Mezi hlavní výhody MC integrování patří

- Jednoduchá implementace,
- Robustní řešení pro různé tvary domén a integrantů,
- Efektivní pro vícerozměrné integrály.

Její nevýhody jsou

- Relativně pomalá konvergence – zmenšení statistické chyby o polovinu vyžaduje zvětšit počet vzorků čtyřikrát,
- Pro syntézu obrazu: obrázek obsahuje šum.

### Náhodné veličiny

Mějme pravděpodobnostní prostor  $(\Omega, \mathcal{A}, P)$ . Potom náhodná veličina je reálná funkce

$$X : (\Omega, \mathcal{A}) \rightarrow (R^1, \mathcal{B}^1),$$

kde  $R^1$  je množina všech reálných čísel a  $\mathcal{B}^1$  je množina všech borelovských množin v  $R^1$ .

Distribuční funkce  $F_X$  náhodné veličiny  $X$  je funkce

$$F_X(x) = P(X \leq x),$$

kde

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}.$$

Další text již bude méně formální a více intuitivní.

### Diskrétní

Diskrétní náhodné veličiny mají konečný počet různých hodnot, které nabývají s určitou pravděpodobností. Rozložení této pravděpodobnosti lze zapsat pomocí pravděpodobnostní funkce (vzhledem k diskretnosti se jedná pouze o tabulku hodnot jejichž součet se rovná 1).

## Spojité

Pro spojité náhodné veličiny se místo pravděpodobnostní funkce používá hustota pravděpodobnosti  $p(x)$  (probability density function, **pdf**), pro kterou platí

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

Distribuční funkce  $F_X(x)$  (cumulative distribution function, **cdf**) se tedy vypočte jako

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

## Vlastnosti náhodné veličiny

Mezi základní vlastnosti náhodných veličin patří střední hodnota a rozptyl. Mějme náhodnou veličinu  $X$ . Střední hodnotu  $E[X]$  definujeme jako

$$E[X] = \int_{\Omega} xp(x) dx.$$

Rozptyl  $V[X]$  definujeme jako

$$V[X] = E[(X - E[X])^2].$$

Lze ukázat, že

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

Pro nezávislé náhodné veličiny  $X_i$  platí, že

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

Pro  $a \in \mathbb{R}$  platí, že

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

Máme-li funkci  $f$ , tak  $f(X)$  je též náhodná veličina a platí, že

$$E[f(X)] = \int_{\Omega} f(x)p(x) dx.$$

## Estimátory obecně

Estimátor je náhodná veličina, konstruovaná za účelem odhadu nějaké veličiny, v našem případě integrálu dané funkce  $f(x)$ . Estimátor je zpravidla definován jako vhodná transformace jiné náhodné veličiny. Realizací (hodnotou) estimátoru je konkrétní odhad.

### Nestranost estimátoru

Estimátor je nestranný pokud jeho střední hodnota je skutečná odhadovaná hodnota. Je-li náhodná veličina  $F$  estimátor hodnoty  $Q \in \mathbb{R}$ , tak estimátor  $F$  je nestranný, jestliže

$$E[F] = Q.$$

Pokud estimátor  $F$  není nestranný tak hodnotě

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

říkáme výchylka (**bias**) estimátoru  $F$ .

## Konzistence estimátoru

Máme-li estimátor  $G$  tvaru

$$G = \sum_i^N F_i$$

odhadující hodnotu  $Q \in \mathbb{R}$ , kde  $F_i$  jsou nezávislé stejně rozdělené náhodné veličiny, pak  $G$  je konzistentní, jestliže

$$P\left(\lim_{N \rightarrow \infty} G = Q\right) = 1.$$

Postačující podmínka pro konzistenci estimátoru je

$$\lim_{N \rightarrow \infty} V[G] = 0.$$

To znamená, že ne každý nestranný estimátor je konzistentní. Příkladem takového estimátoru může být náhodná veličina s nulovou střední hodnotou a nekonečným rozptylem

### Střední kvadratická chyba estimátoru

Střední kvadratická chyba estimátoru  $F$  hodnoty  $Q \in \mathbb{R}$  je definovaná jako

$$MSE[F] = E[(F - Q)^2].$$

Z definice střední hodnoty lze snadno ukázat, že platí

$$MSE[F] = V[F] + \beta[F]^2.$$

Pokud je  $F$  nestranný, tak  $\beta[F] = 0$  a  $MSE[F]$  můžeme vypočítat jako rozptyl  $F$ . Pokud je  $F$  stejného tvaru jako v případě estimátoru, u kterého jsme definovali konzistentnost, tak  $MSE[F]$  můžeme odhadnout nestranným výběrovým rozptylem z jednotlivých vzorků.

### Účinnost estimátoru

Chceme-li pro nestranný estimátor  $F$  zachytit míru účinnosti, která bere v potaz dobu výpočtu a jeho chybu, tak můžeme učinit například následovně

$$\epsilon[F] = \frac{1}{V[F]T[F]},$$

kde  $T[F]$  je čas výpočtu daného estimátoru.

## Odhad integrálu metodou Monte Carlo

Budeme odhadovat integrál  $I$  funkce  $f$  definovaný jako

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

### Primární estimátor

Primární estimátor  $F_{\text{prim}}$  integrálu  $I$  definujeme vzorcem

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

kde  $X$  je libovolná náhodná veličina s hustotou pravděpodobnosti  $p(x)$ , pro kterou platí

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

## Nestrannost primárního estimátoru

Nestrannost primárního estimátoru plyne z jeho definice

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

## Rozptyl primárního estimátoru

Rozptyl se používá jako měřítko pro určení chyby estimátoru.

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

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 odhadu, 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:

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

Z odvození nestrannosti vidíme, že hustota pravděpodobnosti obecně nemusí být identická pro všechny vzorky.

## Rozptyl sekundárního estimátoru

Pro rozptyl sekundárního estimátoru platí vztah

$$V[F_N] = V\left[\frac{1}{N} \sum_i Y_i\right] = \frac{1}{N^2} NV[Y_i] = \frac{1}{N} V[F_{\text{prim}}].$$

Rozptyl je tedy přímo úměrný rozptylu primárního estimátoru s koeficientem  $1/N$ . 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 umístí uje vzorky přednostně do takových oblastí. Toho se docílí tím, že pdf ze které se vybírají vzorky bude podobná integrandu. Použitím vzorkování podle 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 integrál  $p(x)$  přes celou doménu je 1 a tedy, že  $p(x)$  je hustota pravděpodobnosti. Rozptyl by pak byl nulový, jak je vidět z následujících úprav.

$$\begin{aligned} V[F_{\text{prim}}] &= \int_{\Omega} \frac{f(x)^2}{p(x)} dx - I^2 = \int_{\Omega} f(x) \cdot N dx - I^2 = \\ &= N \int_{\Omega} f(x) dx - I^2 = \left( \int_{\Omega} f(x) dx \right)^2 - I^2 = 0 \end{aligned}$$

Bohužel v praxi se taková pdf nedá použít, protože pro její konstrukci bychom potřebovali znát hodnotu integrálu, který se snažíme spočítat.

## Odhad integrálu pomocí řídicí funkce (Control Variate)

Jednou z možností jak snížit rozptyl odhadu je využití takzvané řídicí funkce  $g$ . Důležité je aby tato funkce byla analyticky integrovatelná a zároveň aproximovala námi integrovanou funkci  $f$ .

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

Pravý člen posledního výrazu umíme zintegrovat analyticky a levý člen integrujeme numericky jako doposud. Výhodou je to, že čím jak funkce  $g$  aproximuje funkci  $f$ , tak jejich rozdíl bude mít menší rozptyl. Tato metoda je lépe použitelná v případě, kdy se analyticky integrovatelná funkce vyskytuje v integrované funkci jako aditivní člen. To v syntéze obrazu bohužel neplatí a využívá se proto téměř vždy importance sampling.

## Vzorkování po částech (stratified sampling)

Při hledání vzorků pro odhad integrálu dochází v popsaných metodách často k náhodnému shlukování prvků. Tyto shluky silně přispívají k velikosti rozptylu. Popíšeme dvě metody snažící se o potlačení těchto shluků. První z nich je právě vzorkování po částech a jak název napovídá, tak prostor, přes který integrujeme, rozdělíme do disjunktních částí a odhad integrálu budeme počítat na těchto podmnožinách. Výsledkem bude součet odhadů integrálů na těchto podmnožinách. Tento přístup potlačuje shlukování vzorků a dá se ukázat, že rozptyl takového odhadu bude menší nebo roven rozptylu sekundárního estimátoru se stejným počtem vzorků. Tento postup je velmi účinný pro nízkou dimenzi prostoru, v kterém integrujeme.

Nejjednodušší možností jak prostor rozdělít je uniformní rozklad. Lepších výsledků lze ale dosáhnout, pokud budeme dělení volit tak, aby rozptyl na jednotlivých částech prostoru byl co nejmenší.

## Metody Quasi Monte Carlo

Místo náhodného vzorkování se používají čistě deterministické metody volby vzorků. Prezentované výsledky platí i pro QMC metody, ale v důkazech nelze využít statistiky. Determinismem se snažíme odstranit některé vlastnosti náhodných vzorků, které nám vadily a to především shlukování. Z toho důvodu zavádíme pojem diskrepance.

Nejdříve mějme funkci

$$L(z) = \begin{cases} 1 & \text{pro } 0 \leq z_1, \dots, 0 \leq z_s \\ 0 & \text{jinak,} \end{cases}$$

kde  $z \in \mathbb{R}^s$ . Objem kvádrů definovaných funkcí  $L$  v  $A \subseteq \mathbb{R}^s$  je

$$V(A) = \prod_{j=1}^s v_j.$$

Víme, že Monte Carlo odhad objemu tohoto kvádrů je roven

$$\frac{1}{N} \sum_{i=1}^N L(z_i) = \frac{m(A)}{N},$$

kde  $N$  je počet vzorků,  $z_i$  jsou jednotlivé vzorky a  $m(A)$  značí počet vzorků, které spadly kvádrů  $L$ . Diskrepance množiny bodů je pak definovaná jako maximální možná chyba odhadu objemu přes všechny možné kvádrů v daném prostoru. Tedy diskrepance je

$$D^*(z_1, \dots, z_N) = \sup_A \left| \frac{m(A)}{N} - V(A) \right|.$$

Diskrepance slouží jako míra uniformity dané množiny. Konverguje k nule pro  $N \rightarrow \infty$ .

Existuje teoreticky odhad horní hranice chyby MC odhadu integrálu funkce, který je omezený součinem právě diskrepance a variací integrované funkce. To je důvod proč se snažíme najít vzorkovací sekvence s co nejnižší diskrepancí. Jedna ze sekvencí s nízkou diskrepancí je Van der Corputova a její rozšíření do prostorů s vyšší dimenzí od Haltona nebo Hammersleyho.

## Odhad irradiance

Nyní již zmíníme pouze několik příkladů využití Monte Carlo integrace v počítačové grafice. 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 výsledek 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 nevadí, 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 obecnou BRDF

V tomto případě nás nezajímá irradiance, ale odražená radianci. 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 z 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

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

## Reference

OETIKER, T., PARTL, H., HYNA, I., AND SCHLEGL, E., 2011. The not so short introduction to  $\LaTeX$  2 $\epsilon$ .