

Počítačová grafika III (NPGR010)

Přednáška 4: Zobrazovací rovnice

Zápisky z přednášky Jaroslava Krivánka, rozšířeno z materiálů Adama Domince

Jan Navrátil
27. října, 2012

1 Opakování

1.1 BRDF

Bidirectional reflectance distribution function (BRDF) popisuje hustotu pravděpodobnosti, že se foton dopadající na materiál z jednoho určitého směru odrazí do jiného určitého směru. Charakterizuje tedy odraz světla při dopadu na určitý povrch – vyjma jevů, kdy se světlo nějakým způsobem šíří uvnitř tělesa.

Protože se jedná o hustotu pravděpodobnosti, nejsou její hodnoty nic přímo měřitelného. Pravděpodobnost odrazu do přesně zadaného směru je (ve většině případů, vyjma ideálních zrcadel) nulová; musíme si vypomocet derivací. Půjde o radianci (zář) odraženou určitým směrem, pokud jediný paprsek přichází z diferenciálního – limitně malého – prostorového úhlu. BRDF ve svém celém tvaru tedy vypadá takto:

$$f_r(x, \omega_i \rightarrow \omega_o) = \frac{dL_r(x, \omega_o)}{L_i(x, \omega_i) \cdot \cos(\theta_i) d\omega_i}$$

Zde ω_i vyjadřuje směr dopadajícího paprsku, ω_o směr odraženého paprsku a x je bod na ploše, o kterou se paprsek má odrazet. Hodnota θ_i je velikost úhlu od dopadajícího paprsku k normále plochy a $d\omega_i$ má význam diferenciálního prostorového úhlu. Veličiny L_r , resp. L_i jsou odražená, resp. přicházející radiance. Operátor \rightarrow znamená přechod od přicházejícího úhlu k odchozímu.

1.2 Rovnice odrazu

Ve scénách podle našeho fyzikálního modelu je světlo vyzařováno ze zdrojů, nějakým způsobem prochází scénou a pokud není pohlceno, nakonec dopadne do kamery. Můžeme předpokládat, že je tok světla scénou nějakým způsobem ustálený, tedy na každém bodu x plochy scény máme dobře definovanou funkci $L_i(x, \omega_i)$ radiance přicházející ze směru ω_i a funkci $L_r(x, \omega_o)$ celkové odražené radiance ve směru ω_o . Tyto funkce jsme už jednou svázali přes BRDF. Když zohledníme všechny dopadající paprsky z celé hemisféry H viditelné z bodu x (světlo přicházející zevnitř materiálu neuvažujeme), můžeme z definice BRDF odvodit vztah:

$$L_r(x, \omega_o) = \int_H L_i(x, \omega_i) \cdot f_r(x, \omega_i \rightarrow \omega_o) \cdot \cos(\theta_i) d\omega_i$$

Tento vztah ještě nestačí, protože BRDF na zdrojích světla nezohledňuje přímo vyzařovanou radianci: L_r je jen ta odražená. Můžeme si zavést funkci vyzařované radiance $L_e(x, \omega_o)$, která bude pro každý zdroj pevně daná, a funkci celkové vycházející radiance $L_o(x, \omega_o)$. Protože transport světla splňuje zákon superpozice, můžeme zavést $L_o = L_e + L_r$ a dosazením získáme:

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

Výsledný vztah nám popisuje odchozí radianci z jednoho místa ve scéně a nazýváme jej *rovnici odrazu* (ikdyž zahrnuje víc, než odraz).

2 Zobrazovací rovnice

ZR je rovnice podobného tvaru jako rovnice odrazu, která ale neuvažuje odraženou radianci jako funkci v místě ($L_i(x, \omega_i)$), ale jako produkt působení zbytku scény. Konkrétní podoba rovnice závisí na způsobu, jak scénu chápeme - jako množinu směrů, odkud radiance přichází, nebo jako množinu elementárních plošek, které radianci posílají.

2.1 Úhlový tvar

Všimneme si, že přicházející radiance L_i nutně vychází z některé jiné plochy scény. Zavedeme si tedy funkci *vržení paprsku* $r(x, \omega)$ (raycast) z bodu x ve směru ω , která navrací opět bod scény – právě ten, který je ve směru ω vidět. Paprsek přicházející ze směru ω_i může vycházet jediné z bodu $r(x, \omega_i)$ ve směru opačném, tedy $-\omega_i$. Protože se radiance podél paprsku nemění, získáváme vztah $L_i(x, \omega_i) = L_o(r(x, \omega_i), -\omega_i)$. Dosazením pravé strany této rovnice do rovnice odrazu získáváme úplný matematický popis našeho fyzikálního modelu:

$$L_o(x, \omega_o) = L_e(x, \omega_o) + \int_{\omega_i \in H(x)} L_o(r(x, \omega_i), -\omega_i) \cdot f_r(x, \omega_i \rightarrow \omega_o) \cdot \cos(\theta_i) d\omega_i$$

Místo L_o můžeme klidně psát jen L , protože v rovnici se už vyskytuje pouze odchozí radiance L_o .

Aby scéna byla z našeho hlediska v pořádku, musí tento vztah splňovat. Reálné scény jej splňují vždy, totiž pokud zanedbáme šíření světla pod povrchem a další jevy, na které náš fyzikální model nestačí.

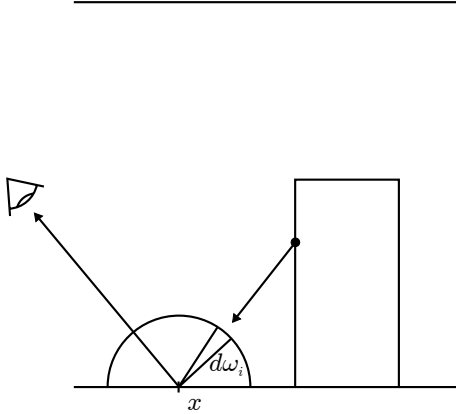
Vztah se nazývá *zobrazovací rovnice* a formuloval jej (v trochu odlišné podobě) James T. Kajiya ve svém článku *The Rendering Equation* v roce 1986. Má význam mimo jiné proto, že mnoho přístupů k vykreslování jsou jen hrubými přístupy řešení této rovnice s tím, že zanedbáme některé (občas velmi významné) faktory.

Například, obyčejné vykreslování na způsob OpenGL bere v potaz jen přímé osvětlení, takže dopadající radianci integruje jen přes světelné zdroje. Modely BRDF používané v OpenGL bývají také daleko od fyzikální reality, ale to je spíše věc volby. Raytracing pak tento pohled rozšiřuje o to, že zohledňuje nepřímé osvětlení na zrcadlově lesklých plochách.

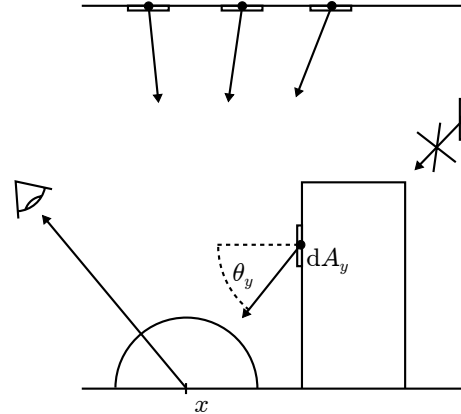
2.2 Plošný tvar zobrazovací rovnice

Jiný přístup má *radiační metoda* vykreslování (*radiosity*). Než se jí začneme věnovat, pro přehlednost si přepíšeme (substituujeme) zobrazovací rovnici do jiného tvaru, ve kterém vystupuje integrál

Obrázek 1: Integrace podle úhlu



Obrázek 2: Integrace přes plochu scény



přes celou plochu scény M (viz obr. 1, 2). Zavedeme si pak navíc radianci z bodu y do bodu x jako $L(y \rightarrow x)$. To bude přirozeně radiance vycházející z y směrem do x , tedy $L(y \rightarrow x) = L(y, \frac{x-y}{\|x-y\|})$. Podobným způsobem zavedeme BRDF $f_r(y \rightarrow x \rightarrow \omega_o)$ pro paprsek přicházející ze směru daného bodem y . Celá zobrazovací rovnice je pak zcela stejně platná v následujícím plošném tvaru:

$$L(x, \omega_o) = L_e(x, \omega_o) +$$

$$\int_M L(y \rightarrow x) \cdot f_r(y \rightarrow x \rightarrow \omega_o) \cdot G(x \leftrightarrow y) \cdot V(x \leftrightarrow y) dA_y$$

Integrujeme přes všechny body y na ploše scény M , nebo přesněji přes dA_y , jejich limitně malou (diferenciální) okolní plochu. V úhlovém tvaru zobrazovací rovnice jsme ale integrovali jen přes body, které dokázala navrátit funkce vržení paprsku $r(x, \omega)$ pro nějaké ω z viditelné hemisféry. Ostatní body musíme z integrálu explicitně odstranit, což zajistí činitel $V(x \leftrightarrow y)$ – viditelnostní funkce. Ta nabývá hodnoty 1, pokud mezi body x a y může vést přímý paprsek, a hodnoty 0, pokud je některá z ploch natočená odvráceně nebo mezi nimi překáží nějaká další plocha.

Dále z původního úhlového integrálu zůstane geometrický člen $G(x \leftrightarrow y)$ zohledňující vzájemnou polohu obou bodů a natočení jejich ploch:

$$G(x \leftrightarrow y) = \frac{\cos(\theta_x) \cdot \cos(\theta_y)}{\|x - y\|^2}.$$

Vyplyvá z členu $\cos(\theta_x) \equiv \cos(\theta_i)$ z původního integrálu a dále ze substituce za $d\omega_i$ podle vztahu diferenciální plošky a diferenciálního prostorového úhlu. Ve výše uvedené substituci konkrétně platí

$$d\omega = dA \cdot \cos(\theta) \cdot r^{-2}.$$

Mimochodem, v plošném tvaru se zobrazovací rovnice blíže podobá Kajiyově původní formulaci.

2.3 Radiozita

Pohled radiální metody do rovnice vnese mnoho zjednodušujících předpokladů. Nejvýznamnější je, že všechny plochy jsou lambertovsky difúzní, tedy že jejich BRDF je ve všech směrech konstantní. V přesném fyzikálním zápisu vychází z odrazivosti: $f_r(x, \omega_i \rightarrow \omega_o) = \frac{1}{\pi} \rho(x)$. Dále místo radiance $L(y \rightarrow x)$ můžeme použít intenzitu vyzařování $B(y)$, která (právě že) není závislá na směru. Za těchto předpokladů můžeme zobrazovací rovnici zapsat ve tvaru:

$$B(x) = B_e(x) + \int_M \frac{1}{\pi} \rho(x) B(y) \cdot G(x \leftrightarrow y) \cdot V(x \leftrightarrow y) dA_y$$

Je vidět, že některé činitele můžeme vytknout z integrálu ven. Když navíc zjednodušíme svoje představy tak, že scénu rozdělíme na konečný počet plošek $M_1 \dots M_N$ s konstantní intenzitou vyzařování $B_1 \dots B_N$ a odrazivostí $\rho_1 \dots \rho_N$, můžeme si rovnici přepsat tak, aby pracovala jen s přenosem světla mezi dvojicemi plošek.

Aby byla B_j v každé plošce konstantní, budeme její hodnoty průměrovat. To je opět velmi hrubý odhad: pokud je například některá ploška zčásti zastíněná, rozloží se do její plochy rovnoměrně (zde vyvstává otázka dostatečně jemné teselace). Dostáváme se díky tomu ale na soustavu s konečným počtem neznámých $B_1 \dots B_N$, kterou můžeme stručně popsat sumou:

$$B_j = B_{e_j} + \frac{1}{\pi} \rho_j \sum_{k=1}^N B_k \cdot F_{k \rightarrow j}$$

Veličina B_{e_j} je intenzita vyzařování zdroje, kterou odhadneme také jako konstantní po celé plošce a pro všechny směry.

Integrál nemohl úplně zmizet – jen se schoval do činitele $F_{k \rightarrow j}$. V něm se podrobně projeví vzájemná poloha obou plošek, jejich natočení a viditelnost mezi nimi (můžou být zakryté třeba jen zčásti). Má vyjadřovat právě, do jaké míry intenzita vyzařování plošky M_k ovlivňuje plošku M_j . Bude ve tvaru:

$$F_{k \rightarrow j} = \frac{1}{|M_j| \cdot \pi} \iint_{M_k M_j} G(x \leftrightarrow y) \cdot V(x \leftrightarrow y) dA_y dA_x$$

Jeden integrál dokonce přibyl, ale to je v pořádku: zlomek $\frac{1}{|M_j|}$ nám připomíná, že průměrujeme hodnoty přes celou plošku M_j , jejíž plochu pochopitelně označujeme $|M_j|$. Integrujeme přes všechny body y uvnitř plošky M_k a přes všechny body x uvnitř M_j (přesněji přes jim příslušná limitní okolí dA_y, dA_x).

Hodnotě $F_{k \rightarrow j}$ se říká *konfigurační faktor* nebo též *form factor*. Musíme jej nějakým způsobem spočítat před samotným vykreslováním – do určité míry můžeme integrovat analyticky, ale (předešlým kvůli viditelnostní funkci) bude lepší jej nějak numericky

odhadnout. Už kvůli rychlosti, protože celkem ve scéně potřebujeme spočítat $O(N^2)$ konfiguračních faktorů. Další problém radiční metody je, že pro protínající se plochy nemůže konfigurační faktor vyjít nijak smysluplně – v takových oblastech se opravdu vykreslí chyby.

Potom už je výpočet snadný: rozepíšeme sumu v rovnicích pro všechna B_j do matice a vyřešíme tuto lineární soustavu. Bude mít N^2 prvků, takže i nepřesná aproximace může trvat dlouho. Výsledné intenzity vyzařování můžeme přímo použít pro vykreslování scény. Aby obrázek nebyl tolik zubatý, obvykle jejich hodnoty interpolujeme mezi sousedními ploškami.

2.4 Vyjádření ZR pomocí integrálního operátoru

Vrátíme se k řešení zobrazovací rovnice. Formálně vzato se jedná o integrální rovnici druhého druhu, která v obecném tvaru vypadá takto:

$$f(x) = g(x) + \int k(x, t) \cdot f(t) dt$$

Tohle zjištění nám ale mnoho nepomůže, spíš naopak – můžeme si být jisti, že ji neumíme vyřešit analyticky. Nezbyvá než zkusit její řešení odhadnout.

Pro stručnější zápis integrálu si zavedeme operátor T , též *transportní operátor*. Chceme jej použít na funkci vycházející radiance $L(x, \omega)$, a proto tento operátor bude zobrazování z množiny funkcí typu $f : M \times S \rightarrow \mathbb{R}$ do té samé množiny (kde S je jednotková koule, tj. všechny směry). Definujeme jej takto:

$$(T \circ L)(x, \omega_o) = \int_H L(r(x, \omega_i), -\omega_i) \cdot f_r(x, \omega_i \rightarrow \omega_o) \cdot \cos(\theta_i) d\omega_i$$

Smysl je jasný: můžeme jej dosadit do zobrazovací rovnice (v jejím úhlovém tvaru) a její zápis se tím působivě zjednoduší. Následující vztah má platit na celém definičním oboru:

$$L = L_e + T \circ L$$

Z algebraického úhlu pohledu také vidíme, že T je lineární operátor. Podrobněji to znamená, že pro konstanty a, b a funkce f, g vhodného typu platí vztah $T \circ (a \cdot f + b \cdot g) = a \cdot (T \circ f) + b \cdot (T \circ g)$. Ověřit tuto vlastnost lze přímo dosazením do integrálu.

Z intuitivního pohledu zastupuje T jeden odraz světla od stěn scény. Například, pokud jej použijeme na funkci vyzařování zdrojů L_e , v každém bodě scény se tím zintegruje dopadající zář, vážená přes BRDF – proto $T \circ L_e$ vyjadřuje přímé osvětlení scény, tj. světelné cesty, které se na cestě od zdroje do kamery odrazí právě jednou. Podobně, $T \circ T \circ L_e$ vyjadřuje osvětlení (výhradně) pomocí světelných cest, které se na cestě od zdroje odrazily právě dvakrát atd.

Mimoходом, na problém můžeme nahlížet tak, že rovnici upravíme jako $L = (I - T)^{-1} \circ L_e$. Výsledek inverze $(I - T)^{-1}$ je pak další operátor nad prostorem funkcí typu $M \times S \rightarrow \mathbb{R}$. Bohužel jej neznáme – dává právě řešení, které se snažíme aproximovat.

2.5 Zápis zobrazovací rovnice řadou

Zobrazovací rovnici můžeme samu do sebe (rekurzivně) substituuovat. Má to smysl – jednou substitucí získáme $L = L_e + T \circ (L_e + T \circ L)$ a protože T je lineární, také $L = L_e + T \circ L_e + T^2 \circ L$. Snadno

vidíme, že n -násobným opakováním substituce vznikne Neumannova řada v tomto tvaru:

$$L = \sum_{i=0}^n (T^i \circ L_e) + T^{n+1} \circ L$$

Pro dokonalou simulaci v rámci modelu $n \rightarrow \infty$. Velmi důležité pozorování je, že T je kontrakce, formálně řečeno $\|T \circ L\| \leq \|L\|$. Míjíme tím, že celkové množství světla se aplikací operátoru zmenší. To plyne přímo ze zákona zachování energie – odražené paprsky můžou mít v součtu nejvýše stejnou energii jako ty dopadající. (V praxi se dokonce vždy nějaký podíl dopadajícího světla v materiálu pohltí, běžné materiály odráží nejvýše asi $4/5$ dopadající světelné energie.)

Protože je T kontrakce, můžeme jeho vysoké mocniny zanedbat – pro n vyšší než nějaké n_0 (čím lesklejší scéna, tím větší n_0) bude příspěvek $T^{n+1} \circ L$ velmi malý. Formálněji, platí: $\lim_{n \rightarrow \infty} (T^n \circ L) = 0$. Díky tomu má smysl opakovat substituci limitně do nekonečna, kdy se člen $T^{n+1} \circ L$ docela ztratí a získáváme nekonečnou řadu:

$$L = \sum_{i=0}^{\infty} T^i \circ L_e$$

Obecně se řada v tomto tvaru s i -tou mocninou operátoru nazývá *Neumannova řada* (po pruském matematikovi Carlu Neumannovi).

To, co nám zůstalo, už není integrální rovnice, ale jen konvergentní řada obsahující integrály (postupně rostoucí dimenze (T^1 je dvou-rozměrný integrál, T^2 čtyřrozměrný atd.). To je mnohem lepší: když vyčíslíme dostatečné množství jejích členů, správnému řešení se můžeme libovolně přiblížit.

2.6 Integrace

Zbývá tedy nějakou numerickou metodou odhadnout danou řadu integrálů. Potíž je jednak v tom, že integrand je významným způsobem nespojitý (scéna obvykle obsahuje mnoho ostrých hran, složitě textury apod.) – obvyklé kvadraturní metody (Gaussova, Rombergova) kvůli tomu nedosahují nijak zvlášť přesných výsledků. Hlavní problém ale je, že chceme počítat s vícenásobnými integrály (každá mocnina operátoru T znamená další integraci přes dva rozměry směru). Téměř všechny kvadraturní vzorce mají společné, že počet potřebných vzorků roste exponenciálně s počtem rozměrů, pokud se má zachovat určitá přesnost. To je nepřijatelné.

Použijeme metodu *Monte Carlo*, integraci náhodným vzorkováním. Obrovskou výhodou této metody je nezávislost konvergence chyby na dimenzi integrandu. Jinými slovy počet vzorků integrandu potřebných pro dosažení určité přesnosti je zcela nezávislý na počtu rozměrů integrálu. Na N vzorcích počítá MC metoda s chybou $O(\frac{1}{\sqrt{N}})$. To je horší výsledek než dávají běžná kvadraturní vzorky pro jednorozměrné integrály, avšak pro integrály dvou a vícerozměrných funkcí již nemá použití běžných kvadraturních vzorců oproti MC metodám žádnou výhodu. MC metoda dále neklade žádné nároky na chování funkce (jako je spojitost).

2.7 Shrnutí

Pro vykreslení obrázku potřebujeme spočítat hodnoty radiance, která přichází z určitých směrů (daných polohami pixelů) do kamery. Pro zjištění bodů, ze kterých vychází, použijeme funkci vržení paprsku $r(x, \omega)$. V každém pixelu tedy potřebujeme spočítat hodnotu radiance $L(r(x_{kamera}, \omega_{pixel}), -\omega_{pixel})$.

Pro její hodnotu jsme si odvodili pohodlný vzorec, zbývá tedy odhadnout hodnotu integrálů $(T \circ L_e)$, $(T \circ T \circ L_e)$, \dots – přímé osvětlení v tomto bodě, dále osvětlení se vzrůstajícím počtem odrazů. Nejvhodnější se zdá použít metodu Monte Carlo. Pro výpočet prvního integrálu tedy budeme náhodně vrhat z tohoto bodu paprsky do zdrojů, pro ten druhý budeme náhodně zkoušet světelné cesty s jedním odrazem atd. Vhodným zpracováním vzorků a sečtením těchto integrálů získáme hodnotu záře (radiance) v daném pixelu.