M2-Images

notions de réduction de variance

J.C. Iehl

October 24, 2022

résumé des épisodes précédents :

bilan :

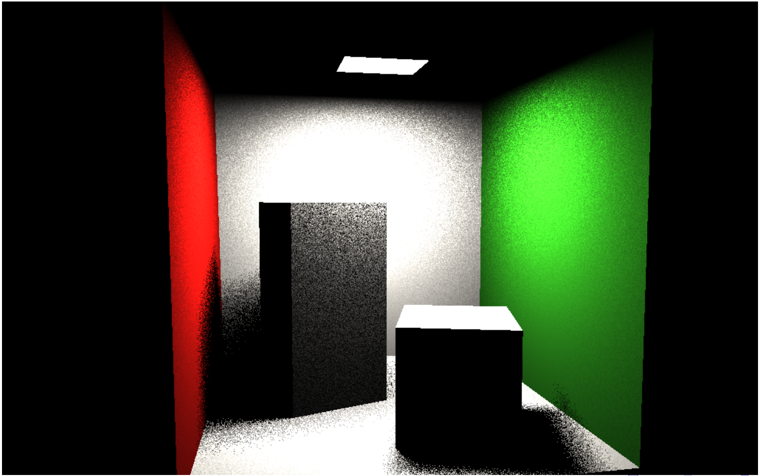
- découper le problème,
- éclairage direct + le reste (*éclairage indirect*)
- viser les sources de lumières,
- sommer la luminance réfléchie, et recommencer...
- l'estimateur *converge* vers la bonne valeur,
- mais il y a du bruit...

le bruit...

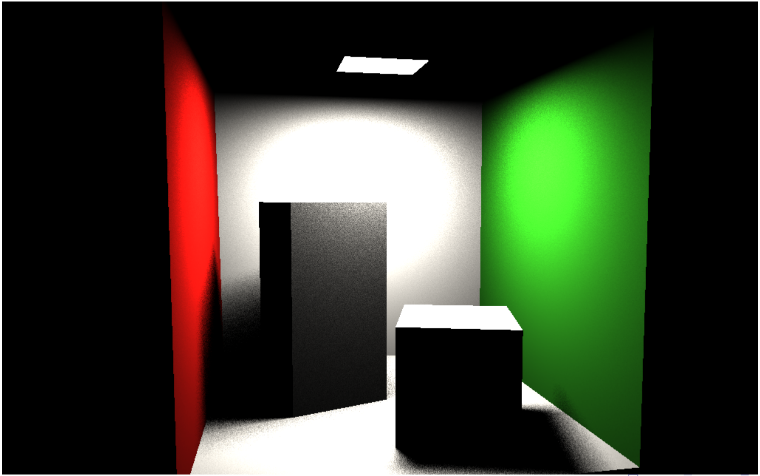
pourquoi ?

- ??
- on veut calculer $I = \int f(x)dx$,
- en utilisant Monte Carlo $I = \frac{1}{N}\sum_i^N \frac{f(x_i)}{p(x_i)}$
- qui converge pour $N = \infty$...
- et pour $N = 4, 16, 64, 256, 1024$?
- il y a des défauts, du bruit...

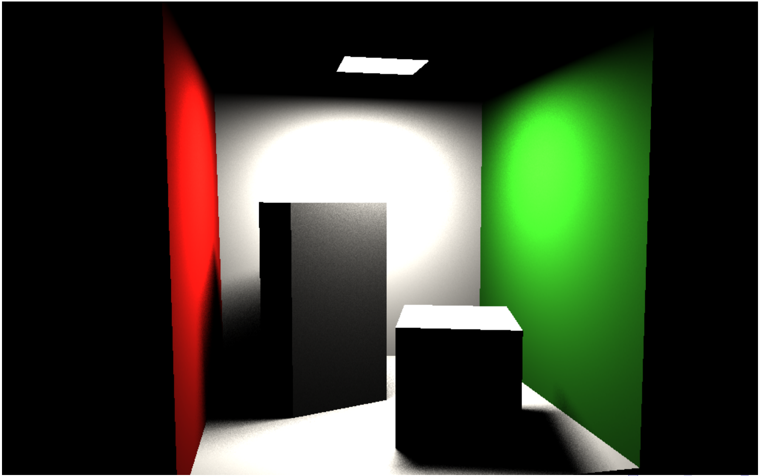
le bruit... 1 point par source



le bruit... 16 points par source



le bruit... 64 points par source



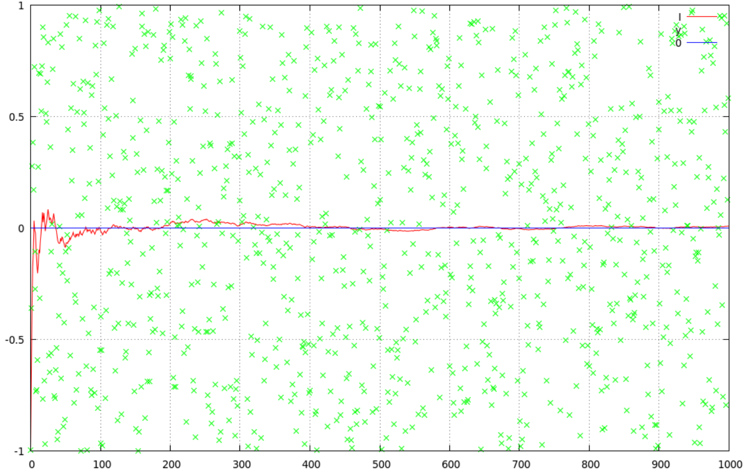
pourquoi ?

- en gros, on calcule la moyenne de N valeurs...

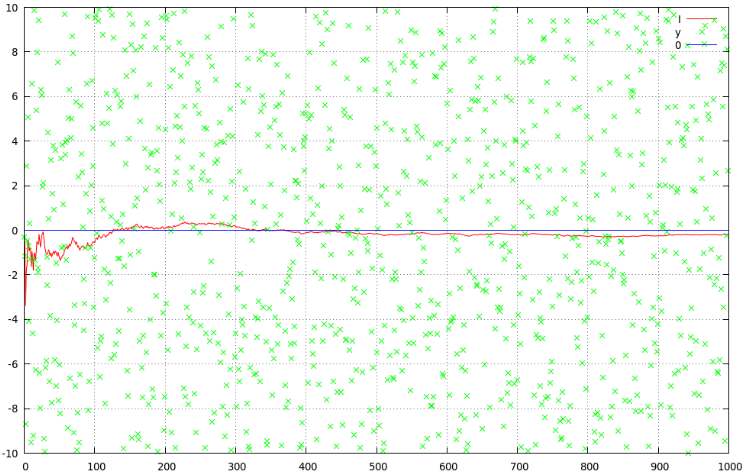
$$I_N = \frac{1}{N} \sum_i^N v_i$$

- comment se comporte le résultat :
- si v_i est une constante ?
- si v_i varie légèrement ?
- si v_i varie fortement ?
- et si ... il y a aussi des accidents ?

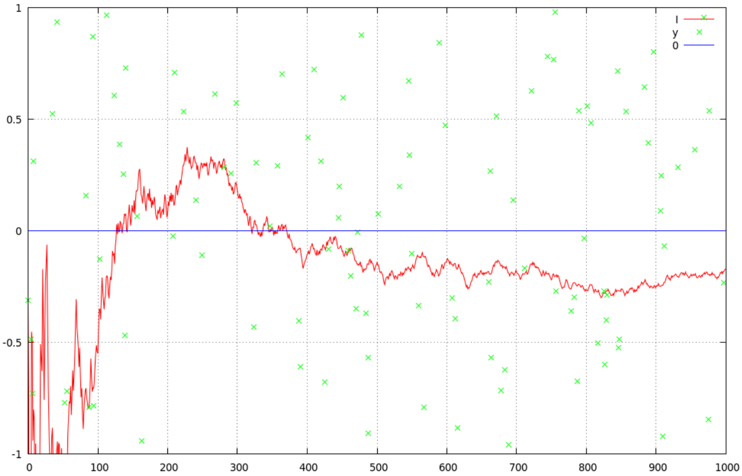
exemple : moyenne de valeurs entre -1 et 1



exemple : moyenne de valeurs entre -10 et 10

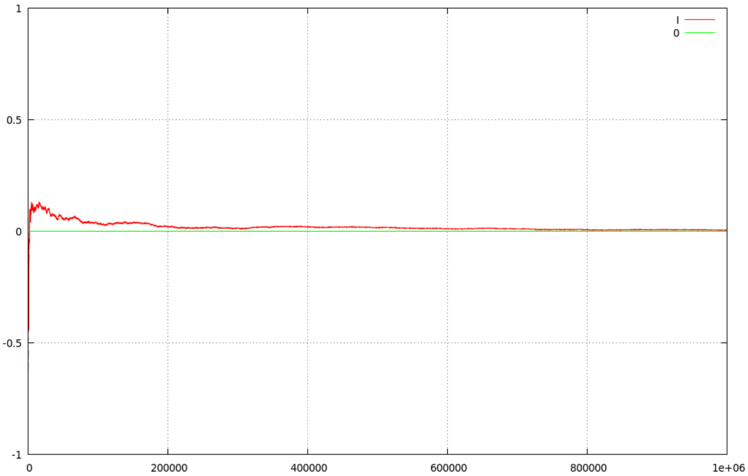


exemple : moyenne de valeurs entre -10 et 10



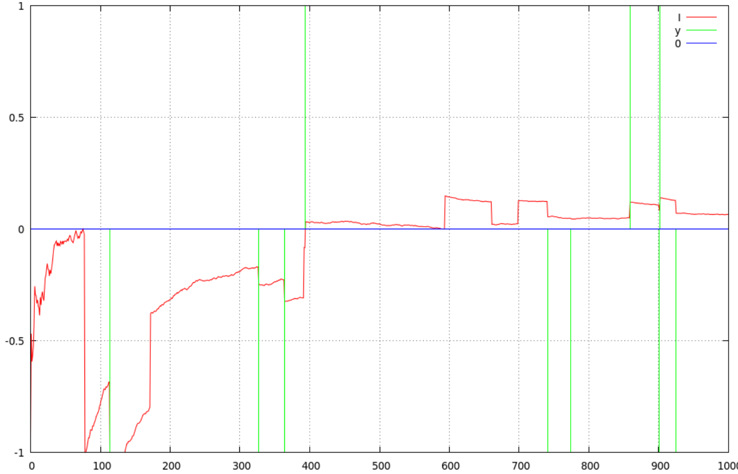
exemple : moyenne de valeurs entre -10 et 10

N = beaucoup plus



exemple : moyenne de valeurs entre -1 et 1

et 1% d'accidents entre -10 et 10



et alors ?

éviter les accidents !

- et les grosses variations de $\frac{f(x_i)}{p(x_i)}$
- choisir le meilleur $p(x)$ possible...
- ??

éviter les accidents ?

lesquels ?

- rappel : (éclairage direct)

$$L_r(x, o) = L_e(x, o) + \int L_e(y, x) f_r(y, x, o) G(x, y) dy$$

$$I \equiv \int L_e(y, x) f_r(y, x, o) G(x, y) dy$$

- et

$$I \approx \frac{1}{N} \sum \frac{L_e(y_i, x) f_r(y_i, x, o) G(x, y_i)}{p(y_i)}$$

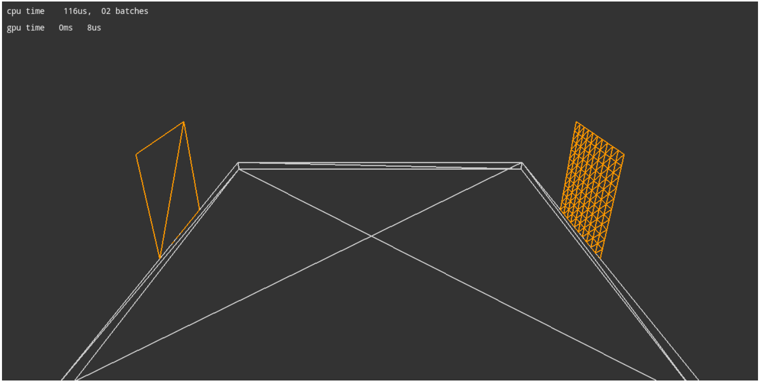
choisir $p(y)$??

réduction de variance

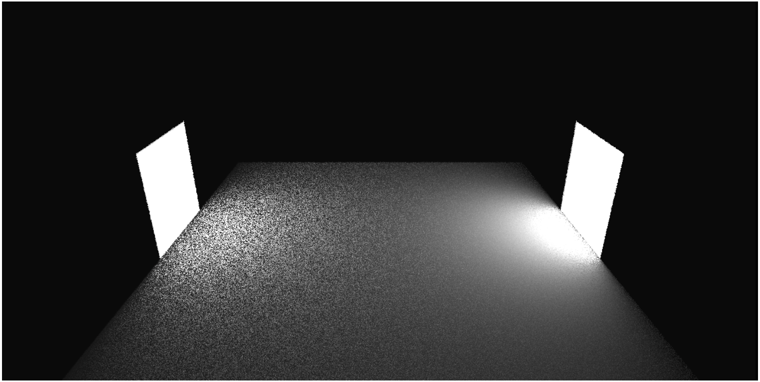
choisir $p(y)$:

- $p(y)$ devrait avoir les mêmes variations que la fonction intégrée...
$$f(y) \equiv L_e(y,x) f_r(y,x,o) G(x,y)$$
- si $f(y)$ à une valeur importante, $p(y)$ doit aussi avoir une valeur importante,
- idem pour les petites valeurs,
- pourquoi ?
- pour limiter les valeurs min et max de $\frac{f(y_i)}{p(y_i)}$!

exemple : éclairage direct



exemple : éclairage direct



$p(y) = \frac{1}{S} \times \frac{1}{Aire(source)}$, avec S le nombre de sources

exemple : $p(y) = \frac{1}{S} \times \frac{1}{Aire(source)}$

pas le même bruit partout ??

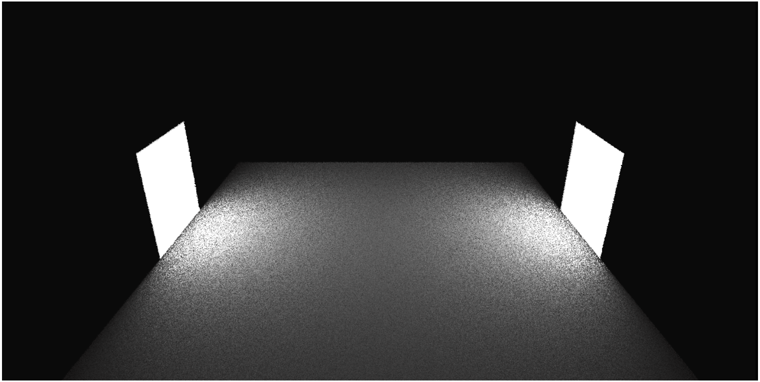
- quelles sont les valeurs de $f(y)$ et de $p(y)$ à gauche de l'image ?
- à droite de l'image ?
- est ce qu'elles varient "correctement" / de la même manière ?
- autre question : combien de points sont choisis sur chaque panneau lumineux ?

chaque panneau mesure 1x1, et le panneau de droite est composé de 100 sources...

on suppose que $L_e(y,x) f_r(y,x,o) V(x,y)$ est constant :

$f(y) = \frac{\cos\theta_x \cos\theta_y}{d^2(x,y)}$ ou encore plus simple : $f(y) = \frac{1}{d^2(x,y)}$

exemple : $p(y) = \frac{1}{\sum_s^S Aire(source_s)}$



$p(y) = \frac{Aire(source)}{\sum_s^S Aire(source_s)} \times \frac{1}{Aire(source)}$, avec S le nombre de sources

exemple : $p(y) = \frac{1}{\sum_s^S Aire(source_s)}$

qu'est ce qui à changé ?

- mêmes questions ?
- ...

exemple : $p(y) = \frac{1}{\sum_s^S Aire(source_s)}$

conclusion :

- ce n'est pas vraiment mieux,
- c'est juste aussi mauvais partout...

$p(y)$ doit inclure le terme $\frac{1}{d^2(x,y)}$ pour faire mieux...

échantillonnage préférentiel

importance sampling

quelle est la densité optimale ?

- on voudrait $\frac{f(y)}{p(y)} = k$,
- pour obtenir la meillleure convergence possible...
- pourquoi n'est elle pas utilisable ?

peut on utiliser n'importe quelle fonction comme densité $p(y)$?

comment construire une densité à partir d'une fonction ?

rappels : densité de probabilité

quelques propriétés :

- la pdf doit être positive,
- la pdf doit être normalisée : $\int p(y)dy \equiv 1$

$p(y)$ peut être nulle si $f(y)$ est également nulle.

exemple $\cos\theta$

utiliser $\cos\theta$ comme pdf :

- on sait que : $0 < \cos\theta < 1$
  (et si $\cos\theta = 0$ alors f aussi...)
- et
$$\int_\Omega \cos\theta d\omega = \pi$$
- donc :
$$\int_\Omega \frac{\cos\theta}{\pi} d\omega = 1$$

bilan : on peut utiliser $\frac{\cos\theta}{\pi}$ comme pdf...
reste un dernier détail : peut on choisir aléatoirement des directions distribuées selon cette pdf ?

exemple $\cos\theta$

générer des directions $\vec{\omega} \propto \frac{\cos\theta}{\pi}$:

- rappel : inversion de la fonction de répartition,
- cf Global Illumination Compendium, eq 35

pdf optimale

quelle est la densité optimale ?

- quelle est la valeur de $\int_A f(y)dy$?
- il faut calculer k tel que : $\int_A \frac{f(y)}{k}dy = 1$,
- pour la normaliser, et l'utiliser comme pdf,
- mais il faut être capable de générer des points $y \propto f(y)$...

et en général ce n'est pas possible...

pdf approchée

et alors ?

- rappel : $f(y) = L_e(y,x) f_r(y,x,o) V(x,y) \frac{\cos\theta_x \cos\theta_y}{d^2(x,y)}$
- ??
- utiliser $L_e(y,x)$ comme pdf ?
- utiliser $f_r(y,x,o)$ comme pdf ?
- utiliser $V(x,y)$ comme pdf ?
- utiliser $\cos\theta_x$ comme pdf ?
- utiliser $\cos\theta_y$ comme pdf ?
- utiliser $\frac{1}{d^2(x,y)}$ comme pdf ?

un "mélange" de plusieurs ?

exemple : $L_e(y, x)$ et $\cos\theta_x$

- $L_e(y, x)$ est utilisable "loin" des sources,
- mais lorsque x est proche de y, $\frac{1}{d^2(x,y)}$ explose...
- rappel : éviter les accidents...
- idée : on sait aussi faire le calcul avec $\cos\theta_x$...
- est-il possible de calculer de plusieurs manières le même résultat et de choisir la meilleure solution ?

échantillonnage préférentiel

multiple importance sampling / MIS

intuition :

- on peut calculer $I_1 = \frac{1}{N} \sum_i^N f(\vec{\omega}_i) / \frac{\cos \theta_i}{\pi}$,
- et aussi $I_2 = \frac{1}{N} \sum f(y_i) / \frac{1}{\sum Aire(source)}$,
- les 2 convergent vers I...
- mais pas de la même manière,
  selon la situation (proche ou loin des sources)...
- comment choisir le meilleur selon la situation ?

est ce que la moyenne des 2 est un meilleur estimateur ?

échantillonnage préférentiel

multiple importance sampling / MIS

choisir le meilleur estimateur ?

- comment les comparer ?
- comment les "mélanger" ?

"Optimally Combining Sampling Techniques for Monte Carlo Rendering"
E. Veach, L.J. Guibas, 1995

échantillonnage préférentiel

multiple importance sampling / MIS

exemple pour 2 estimateurs :

- chaque estimateur est calculé avec sa pdf,
- chaque pdf génère un échantillon y_i,
- les autres pdf évaluent la densité de y_i,
- et les estimateurs sont pondérés proportionnellement...

exemple : MIS $L_e(y, x)$ et $\cos\theta_x$

pour $i = 1..N$ échantillons :

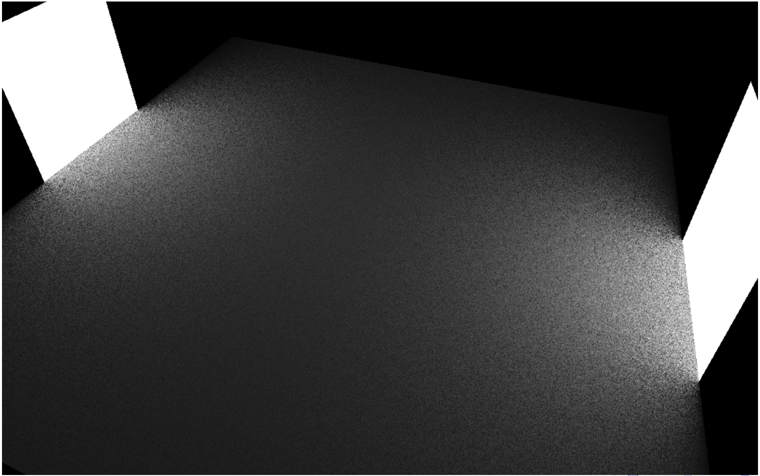
- $p_1(y) = 1/\sum Aire(source)$,
- $p_2(y) = \frac{\cos\theta_x}{\pi} \times \frac{\cos\theta_y}{d^2(x,y)}$ (avec y visible de x dans la direction $\vec{\omega} \propto \frac{\cos\theta_x}{\pi}$)
- générer $y_{1i} \propto p_1()$,
- évaluer $p_2(y_{1i})$,
- générer $y_{2i} \propto p_2()$,
- évaluer $p_1(y_{2i})$
- pondérer les 2 estimateurs :
- $I = \frac{1}{N}\sum_i^N \frac{p_1(y_{1i})}{p_1(y_{1i})+p_2(y_{1i})}\frac{f(y_{1i})}{p_1(y_{1i})} + \frac{1}{N}\sum_i^N \frac{p_2(y_{2i})}{p_1(y_{2i})+p_2(y_{2i})}\frac{f(y_{2i})}{p_2(y_{2i})}$

exemple : MIS $L_e(y, x)$ et $\cos\theta_x$

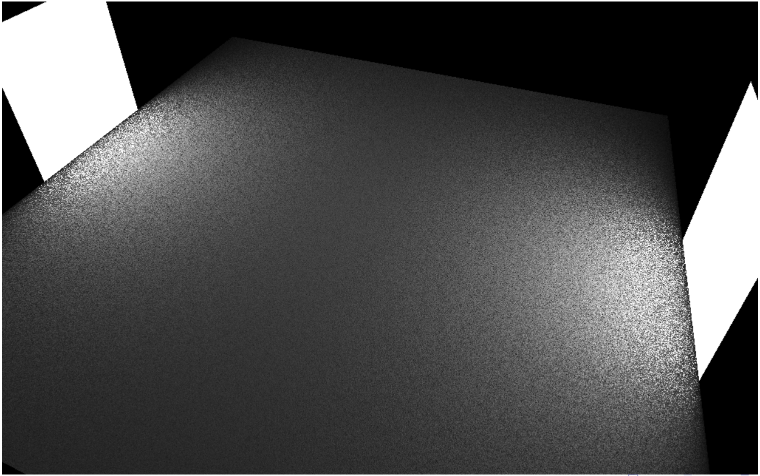
comment ça marche ?

- soit $p_1()$, soit $p_2()$ prend des valeurs importantes en même temps que $f()$,
- et la pondération réduit l'influence des mauvais échantillons / accidents
- d'autres pondérations sont possibles, cf l'article, section 3.4, *power*, *cutoff* et *max* heuristics

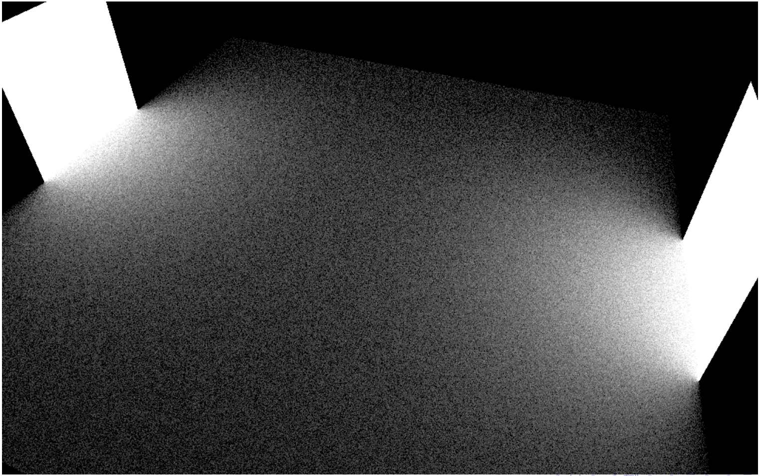
exemple : MIS



exemple : MIS, stratégie 1, $p_1(y) = 1/\sum Aire(source)$



exemple : MIS, stratégie 2, $p_2(y) = \frac{\cos\theta_x}{\pi} \times \frac{\cos\theta_y}{d^2(x,y)}$



exemple : MIS, influence stratégie 2