



Rapport Liens 93-20

Thèse présentée pour obtenir le titre de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

spécialité : informatique

par

Christophe Vedel

**Titre : Calcul de l'illumination par des sources étendues :
espace objet et espace image.**

soutenue le 13 septembre 1993 devant le jury composé de :

MM. Jean-Marc Steyaert
Erik Jansen
Bernard Péroche
Eugene Fiume
Claude Puech
François Sillion

Président
Rapporteur
Rapporteur
Examineur
Examineur
Examineur

Remerciements

Cette thèse a été effectuée au Laboratoire d'Informatique de l'École Normale Supérieure. J'exprime ma gratitude à son directeur, Claude Puech, pour m'y avoir accueilli et avoir dirigé ce travail.

Erik Jansen et Bernard Péroche ont accepté d'examiner ce manuscrit. Qu'ils en soient chaleureusement remerciés.

Je suis également reconnaissant à Eugene Fiume, Jean-Marc Steyaert et François Sillion d'avoir bien voulu juger ce travail.

Mes remerciements vont aussi à tous les membres de l'équipe iMAGIS, Frédéric Asensio, Marc Espie, Marie-Paule et Dominique Gascuel, Nicolas Holzschuch, Alexis Lamouret, Philippe Limantour, François Sillion et Jean-Paul Smets.

Les conseils d'Anne Verroust et de Michel Pocchiola m'ont été d'une aide précieuse dans la dernière étape de ce travail.

Je n'oublie pas Dominique Ho Tin Noé, Annie Iapteff, Fabienne Meunier et Brigitte Van Elsen, ainsi que toute l'équipe du SPI qui à des titres divers m'ont aidé dans l'accomplissement de ce travail.

Introduction

La synthèse d'image réalise de constants progrès dans la modélisation et l'animation d'objets de plus en plus complexes. Parallèlement, les modèles d'illumination qui donnent l'éclairage et donc l'image des objets modélisés ont beaucoup évolué, passant de modèles empiriques à des modèles fondés par la physique, de modèles purement locaux à des modèles globaux, c'est-à-dire intégrant les interréflexions de la lumière dans l'environnement. Il existe cependant un décalage entre modélisation, animation et éclairage. Les algorithmes d'illumination les plus récents comme la radiosité et ses extensions ne s'appliquent pas encore aux multiples formes d'objets que l'on est capable de représenter et le lancer de rayon reste la technique de choix pour beaucoup d'applications.

Plusieurs particularités des nouveaux modèles peuvent expliquer ce retard. Le caractère global des échanges de lumière donne aux algorithmes une complexité quadratique dans la taille de la discrétisation de l'intensité lumineuse sur les surfaces. Des progrès ont été accomplis dans ce domaine avec l'introduction des méthodes de résolution progressive [CCWG88] puis hiérarchiques [HSA91] mais les premiers algorithmes tentant de gérer la complexité de nombreux objets de petite taille commencent seulement à apparaître [RPV93].

Ensuite, la solution de l'équilibre des échanges lumineux rend nécessaire le passage par une représentation de l'intensité lumineuse sur les surfaces dans l'espace objet (maillage). Les méthodes qui ne calculent que l'éclairage direct des lampes n'ont besoin que d'une interface très réduite avec les objets (par exemple, calcul de l'intersection avec un rayon et de la normale au point d'intersection). Par contraste, les méthodes globales sont liées beaucoup plus intimement à la représentation géométrique des objets à laquelle il faut ajouter une représentation lumineuse satisfaisant plusieurs contraintes dont la conservation de l'énergie.

Finalement, l'obtention d'images de bonne qualité requiert une définition précise des ombres importantes qui se traduit par un maillage adaptatif fin des surfaces ou par un recalcul coûteux de l'ombrage direct pour chaque pixel de l'image. Nous explorons dans ce travail ces deux derniers points en présentant un algorithme de maillage adapté aux surfaces courbes et des méthodes d'accélération du recalcul de l'ombrage direct.

Dans le chapitre 1, nous présentons la notion de modèle d'illumination global et la méthode de radiosité qui en est une version simplifiée. Le point de départ

est un modèle local de réflexion de la lumière sur les surfaces qui lie en un point, les flux incident et réfléchi d'énergie lumineuse. En utilisant la conservation du flux dans son transport entre surfaces, on introduit un couplage entre les illuminations de tous les points de la scène et on passe d'un modèle local à un modèle global de l'illumination. Ce couplage peut s'exprimer par une équation intégrale dont la première formalisation dans le domaine graphique est l'*équation du rendu* proposée par Kajiya en 1986. La solution de cette équation est l'objectif formalisé des algorithmes de synthèse d'image réaliste. Dans la suite, nous utilisons une simplification de ce modèle obtenue en ajoutant l'hypothèse de réflexion uniforme de la lumière. Cette approche, la méthode de la radiosité, a fourni les premières images prenant en compte les interreflexions lumineuses dans une scène. Depuis son introduction par Goral en 1984, la méthode a été beaucoup développée et plusieurs algorithmes ont été proposés pour chaque étape du calcul : maillage des surfaces, calcul des facteurs de forme, résolution du système d'équilibre. Nous présentons les principaux qui seront repris et modifiés dans la suite.

Le chapitre 2 passe en revue les méthodes de maillage pour la radiosité. Malgré l'hypothèse simplificatrice sur le mode de réflexion, la fonction de radiosité solution de l'équation intégrale d'équilibre n'est pas calculable exactement. La solution pratique du problème se fait dans un espace de fonctions de dimension finie où l'on cherche une fonction proche de la solution réelle. La multitude des choix possibles pour cet espace : support et forme des fonctions de base, techniques de raffinement adaptatif, etc. a donné lieu à de nombreuses méthodes de maillage. Une des limites de la plupart des méthodes existantes est de n'accepter que des environnements polyédriques. Dans les chapitres qui suivent, nous proposons une technique de maillage apte à traiter les surfaces courbes et utilisant les dérivées de la radiosité pour guider la reconstruction. La première étape consiste à montrer que ces dérivées peuvent être obtenues sans surcoût trop important grâce à une meilleure utilisation des informations données par le calcul des facteurs de forme.

L'étape coûteuse de la méthode de radiosité est le calcul des facteurs de forme entre facettes. En particulier, les tests faits pour déterminer les parties mutuellement visibles des surfaces en interaction fournissent plus d'information que celle qui est réellement utilisée. Dans le chapitre 3, nous montrons comment utiliser cette information supplémentaire pour estimer les dérivées spatiales du vecteur de forme. La variation du facteur est décomposée en deux parties : changement de visibilité de la surface et changement de l'angle solide sous lequel la partie visible est vue depuis le récepteur. La dérivée ainsi obtenue est utilisée par la méthode de maillage décrite au chapitre suivant dans le critère de raffinement adaptatif et lors de la reconstruction.

Le chapitre 4 est consacré à une méthode de maillage qui s'applique aussi bien aux surfaces courbes qu'aux polygones. Une fonction de poids définie dans l'espace est associée à chaque échantillon de radiosité. Lors de la reconstruction, une moyenne pondérée des échantillons proches permet de calculer la radiosité

en tout point. Nous montrons comment améliorer la fonction reconstruite en utilisant les dérivées et, pour les surfaces courbes, des vecteurs de radiosit .

La r solution de l' quation d' quilibre de la lumi re entraine des approximations   toutes les  tapes du calcul : repr sentation de la fonction de radiosit  sur une base finie, calcul approch  des facteurs de forme, r solution it rative du syst me. Ces approximations rendent la pr cision du r sultat final difficile   mesurer. Le chapitre 5 rassemble une s rie de tests visant   estimer cette pr cision. L' clairage sur les surfaces est s par  en une composante directe provenant des sources primaires et une composante indirecte provenant des r flexions diffuses sur l'ensemble des surfaces. Les tests de la premi re partie compare la composante directe obtenue par reconstruction   sa valeur exacte calcul e par int gration de contour. Les points abord s sont l'influence de la subdivision adaptative, de l'utilisation des d riv es pour la reconstruction et du vecteur de forme pour les surfaces courbes. La composante indirecte est plus difficile   tester car sa valeur exacte n'est pas calculable. Pour estimer l' cart   la vraie solution, nous transformons l' quation int grale d' quilibre pour montrer qu'en tout point, la composante indirecte de la radiosit  doit valoir la r flectance au point multipli e par l' clairement ne provenant pas des sources primaires. La technique employ e pour tester la composante indirecte utilise comme mesure, l' cart entre les deux membres de cette  quation, calcul s tous les deux   partir du m me maillage. On examine en particulier la r partition et l'amplitude de l'erreur en fonction de la densit  du maillage, les r sultats d'une technique d' limination des fuites d'ombre propos e au chapitre 4, l'effet des d riv es et des vecteurs de forme dans la reconstruction et l'influence de la finesse de maillage de l' clairage direct sur la composante indirecte.

Le chapitre 5 confirme une importante diff rence entre les composantes directe et indirecte de l' clairage. La composante directe,   cause des ombres marqu es, n cessite un maillage tr s fin pour obtenir une qualit  correcte   la reconstruction. Par contre la composante indirecte varie beaucoup plus lentement et peut  tre calcul e   une bonne pr cision avec relativement peu d' chantillons. Pour obtenir des images de bonne qualit , un choix possible est donc de ne r cup rer dans le maillage que la composante indirecte et de recalculer pixel par pixel et avec la pr cision n cessaire la contribution des lampes. Ce calcul est bien sur consommateur de temps de calcul, en particulier pour des sources de lumi re  tendues. Nous pr sentons dans le chapitre 6 une technique fond e sur le calcul du facteur de forme par int gration de contour. Cette m thode a  t  r serv e jusque l  aux poly dres pour lesquels une formule exacte d'int gration existe. Nous montrons comment l' tendre aux objets quelconques en construisant automatiquement un contour polygonal approch .

Enfin, dans le chapitre 7, nous utilisons les singularit s de la projection des surfaces lisses par morceaux pour faire une classification de leurs discontinuit s d'illumination. Ce r sultat est ensuite appliqu  au calcul de maillages de discontinuit  discrets dans l'espace image pour des surfaces courbes.

Chapitre 1

Modèles d'illumination et radiosit 

1.1 R flexions et interr flexions de la lumi re

Depuis les premi res images de synth se, de nombreux mod les de plus en plus complexes d' clairage des objets ont  t  cr es. Calculant d'abord l'illumination d'un objet de mani re purement locale, ils ont ensuite pris en compte les ombres port es puis les r flexions et transmissions de lumi re entre les objets. Cette  volution a rendu n cessaire l'emploi d'un mod le physiquement correct de la lumi re et de ses interactions avec l'environnement. En effet, les m thodes d'illumination telles que celles de Gouraud ou de Phong se limitent   calculer la contribution directe des lampes et les valeurs prises par l'illumination peuvent donc  tre arbitraires du moment que la correspondance avec l'intensit  des pixels   l' cran donne le r sultat recherch , c'est- -dire un relief et une impression de volume aux objets. Par contre, les mod les plus r cents int gre en chaque point l'illumination provenant de tous les objets de la sc ne et l'existence m me d'une solution d pend alors de propri t  comme la conservation de l' nergie. D'une mani re plus g n rale, le respect des lois de la physique est le moyen le plus s r d'obtenir une simulation r aliste de l' clairage, qui est l'un des objectifs de la recherche dans ce domaine.

La lumi re peut  tre mod lis e plus ou moins finement suivant les ph nom nes qui sont  tudi s : interaction avec la mati re au niveau atomique, interf rences, etc. Pour la synth se d'images, on veut simuler le comportement «macroscopique» de la lumi re et ses effets sensibles   un observateur humain dans les conditions habituelles. La grandeur physique adapt e   cette description est la *luminance* not e $L(x, \vec{\omega})$ qui pour une position x et une direction de propagation $\vec{\omega}$ donn es exprime la puissance lumineuse transport e par unit  de surface et par unit  d'angle solide projet . Avec les notations de la figure 1.1, $L(x, \vec{\omega})$ repr sente la densit  de flux au point x dans la direction $\vec{\omega}$ et le flux   travers l' l ment de

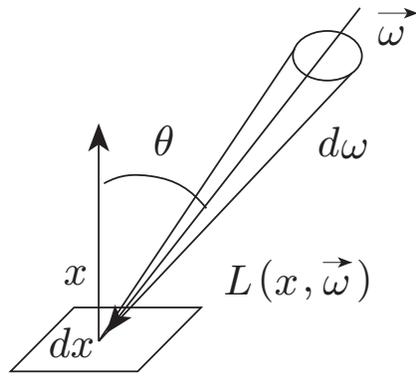


FIG. 1.1 - Géométrie pour la luminance.

surface dx et provenant de $d\omega$ vaut $L(x, \vec{\omega}) \cos \theta dx d\omega$. La luminance se mesure en watts par mètre carré et par stéradian et possède deux propriétés très utiles en pratique :

- en l'absence d'absorption ou de déviation par le milieu traversé, la luminance le long d'un rayon reste constante dans sa propagation ;
- la puissance reçue par l'oeil est proportionnelle à la luminance de la surface regardée.

La première propriété se montre en considérant la conservation du flux lumineux entre deux éléments de surface le long du rayon. La seconde propriété explique que les objets paraissent également éclairés qu'on les regarde de près ou de loin et provient du fait que les capteurs de lumière mesurent le flux total d'énergie incidente. Une autre grandeur importante est l'éclairement ou flux total incident par unité de surface. L'éclairement est l'intégrale de la luminance incidente dans le demi espace situé au dessus de l'élément de surface et se mesure en watts par mètre carré.

Ce choix d'un niveau de description de la lumière adapté aux algorithmes de production d'images ne signifie pas qu'une modélisation plus «microscopique» n'a pas sa place. Un bon exemple est la simulation de la réflexion de la lumière sur des distributions de microfacettes qui permet de construire des modèles de réflexion utilisables ensuite au niveau macroscopique selon une approche habituelle en physique statistique [CT82].

Pour construire un modèle global de l'illumination dans un environnement, on commence par considérer l'interaction locale de la lumière avec une surface, c'est-à-dire le passage d'un flux incident à un flux réfléchi. L'observation que la luminance est conservée dans sa propagation permettra ensuite de coupler les multiples interactions locales [Han92]. En un point x d'une surface, on considère l'énergie lumineuse incidente dans un angle solide $d\omega_i$ (l'éclairement dans cet angle solide). Suivant les caractéristiques du matériau rencontré, une partie de

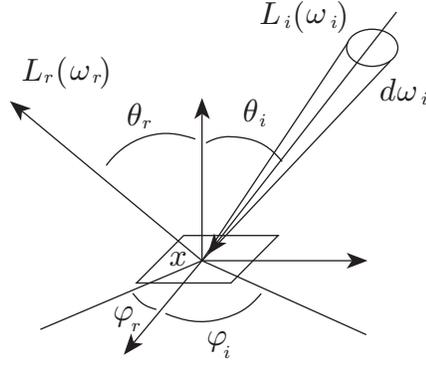


FIG. 1.2 - Géométrie pour la réflexion.

cette énergie est absorbée par la surface et le reste est réémis dans l'hémisphère située au dessus du plan tangent en x . En général, la distribution angulaire de l'énergie réfléchiée dépend de la direction d'incidence ω_i et son amplitude est proportionnelle à l'énergie incidente. Plus précisément, on observe que la luminance réfléchiée dans une direction $\vec{\omega}_r$ est proportionnelle à l'éclairement dans l'angle solide $d\omega_i$ (voir figure 1.2). Le coefficient de proportionnalité dépend des directions $\vec{\omega}_i$ et $\vec{\omega}_r$ et s'appelle la fonction de réflectance bidirectionnelle (ou BRDF en anglais pour *bidirectional reflectance distribution function*). Cette fonction dépend de quatre angles, deux angles polaires θ_i et φ_i paramétrant la direction d'incidence et deux angles polaires θ_r et φ_r paramétrant la direction de réflexion. On la note $f_r(x, \theta_i, \varphi_i, \theta_r, \varphi_r)$ ou $f_r(x, \vec{\omega}_i \rightarrow \vec{\omega}_r)$. Pour la géométrie de la figure 1.2, on peut écrire en intégrant tout l'éclairement :

$$L(x, \vec{\omega}_r) = \int_{\Omega_i} f_r(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) L(x, \vec{\omega}_i) \cos \theta_i d\omega_i \quad (1.1)$$

où Ω_i est l'hémisphère des directions incidentes en x . Ce modèle permet de décrire de nombreux modes de réflexion, par exemple, l'anisotropie est une dépendance dans les angles φ_i et φ_r . Il n'est cependant pas complètement général car il suppose que la réflexion est ponctuelle (pas de réémission en un point différent du point d'incidence) et sans délai (pas de phosphorescence).

A partir de la réflexion locale, on passe à un modèle global en reliant la luminance incidente en x , $L(x, \vec{\omega}_i)$, aux luminances des surfaces visibles depuis x . Soit x' le premier point visible depuis x dans la direction $-\vec{\omega}_i$, on a :

$$L(x, \vec{\omega}_i) = L(x', -\vec{\omega}_i) = L(x', \vec{\omega}_r')$$

Un élément de surface dx' autour de x' est vu depuis x sous l'angle solide $\cos \theta'_r dx'/r^2$, où \vec{r} est le vecteur joignant x à x' . En substituant, $d\omega_i$ par cet angle solide et $L(x, \vec{\omega}_i)$ par $L(x', \vec{\omega}_r')$ dans l'équation 1.1, on obtient :

$$L(x, \vec{\omega}_r) = \int_S f_r(x, \vec{\omega}_i \rightarrow \vec{\omega}_r) L(x', \vec{\omega}_r') \text{vis}(x, x') \frac{\cos \theta_i \cos \theta'_r}{r^2} dx' \quad (1.2)$$

o  S repr sente l'union de toutes les surfaces de la sc ne et $\text{vis}(x, x')$ est la fonction de visibilit  qui vaut 1 si x et x' se voit mutuellement et 0 sinon. Cette fonction permet d' tendre l'int grale   S plut t qu'au sous-ensemble des surfaces visibles depuis x . Dans cette  quation, on note que le terme $(\cos \theta_i \cos \theta'_r)/r^2$ ne d pend que des positions de x et x' puisque θ_i et θ'_r sont les angles entre \vec{r} et les normales en x et x' respectivement. On d finit alors le facteur de forme diff rentiel entre x et x' :

$$G(x, x')dx' = \frac{\cos \theta_i \cos \theta'_r}{\pi r^2} dx',$$

et on obtient finalement la forme int grale de l' quation d' quilibre de la lumi re dans la sc ne :

$$L(x, \vec{\omega}) = L_e(x, \vec{\omega}) + \pi \int_S f_r(x, \vec{\omega}' \rightarrow \vec{\omega}) L(x', \vec{\omega}') \text{vis}(x, x') G(x, x') dx' \quad (1.3)$$

o  on a supprim  les indices r indiquant la r flexion car $\vec{\omega}$ et $\vec{\omega}'$ sont toutes deux des directions r fl chies. $L_e(x, \vec{\omega})$ est la luminance propre  mise au point x pour la direction $\vec{\omega}$. Cette  quation exprime l' quilibre des  changes d' nergie lumineuse entre les objets d'un environnement pour une longueur d'onde de la lumi re donn e. Pour les ph nom nes d' clairage int ressant l'image de synth se, elle est tr s g n rale et recouvre la plupart des mod les existants qui en sont solutions moyennant des approximations plus ou moins grossi res. En particulier, le choix de la fonction de r flectance bidirectionnelle permet de mod liser des types de mat riaux tr s diff rents. Des extensions sont  galement possibles pour la mod lisation de ph nom nes plus complexes comme la travers e de milieux participatifs, constitu s de particules modifiant le trajet de la lumi re dans l'espace entre les objets (poussi re, fum e, vapeur, etc.). Dans le domaine du graphique, la premi re formulation int grale de cet  quilibre, sous une forme l g rement diff rente de celle donn e ici, a  t  propos e par Jim Kajiya en 1986 sous le nom d'* quation du rendu* [Kaj86].

Dans ce travail, nous allons nous int resser essentiellement   la m thode de radiosit . L' quation 1.3 est simplifi e en ajoutant une hypoth se suppl mentaire, la loi de Lambert, qui dit que la r flexion de la lumi re incidente se fait de mani re uniforme dans toutes les directions. La fonction de r flectance bidirectionnelle peut alors s' crire :

$$f_r(x, \vec{\omega}' \rightarrow \vec{\omega}) = \frac{\rho(x)}{\pi}.$$

$\rho(x)$ est la r flectivit  au point x . Pour une surface quelconque, elle est d finie comme le rapport entre le flux total r fl chi et le flux total incident ( clairage). Pour une surface parfaitement diffuse, f_r n'a plus de d pendance angulaire et on a la formule ci-dessus o  le facteur π provient de l'int gration du flux sortant dans l'h misph re. Comme la r mission de l' nergie lumineuse est uniforme, on peut repr senter l' clairage en un point de la surface par une valeur scalaire pour chaque longueur d'onde. La quantit  choisie pour cela est la *radiosit *, not e B ,

c'est-à-dire le flux sortant total en un point, intégrale de la luminance réémise dans l'hémisphère au dessus de la surface. Tenant compte de ces simplifications, on réécrit l'équation 1.3 en :

$$B(x) = B_e(x) + \rho(x) \int_S B(x') \text{vis}(x, x') G(x, x') dx' \quad (1.4)$$

Si on décide en plus d'effectuer un échantillonnage fini des longueurs d'onde du spectre visible, la complexité du problème est grandement réduite, un simple scalaire $B(x)$ remplaçant une fonction de distribution sur l'hémisphère $L(x, \vec{\omega})$. Cependant, la solution de ce problème simplifié reste trop complexe et l'équation intégrale 1.4 n'est pas résoluble de manière symbolique. Il faut donc recourir à des méthodes numériques et chercher une solution discrétisée approchée $\tilde{B}(x)$ de la solution réelle $B(x)$. Pour cela, on commence par choisir un espace de fonctions \mathcal{F} de dimension finie muni d'une base f_i ce qui permet de représenter en machine chaque fonction $\tilde{B}(x)$ par le vecteur (\tilde{B}_i) de ses coordonnées dans la base $\tilde{B}(x) = \sum_i \tilde{B}_i f_i(x)$. On utilise, par exemple, une division en facettes des objets de la scène et des fonctions polynômiales supportées par ces facettes. On peut alors réécrire l'équation intégrale à résoudre

$$\tilde{B}(x) = B_e(x) + \rho(x) \int_S \tilde{B}(x') G(x, x') dx' \quad (1.5)$$

mais cette équation n'admet pas de solution exacte dans \mathcal{F} . Pour minimiser l'écart entre la solution approchée et la vraie radiosit , on utilise l' galit  1.5 pour d finir une fonctionnelle d'erreur que l'on va minimiser en l'annulant autant de fois qu'il y a de fonctions de bases. On obtient ainsi un syst me lin aire dont les inconnues sont les \tilde{B}_i . Les  tapes suivantes du calcul consistent   calculer les coefficients de ce syst me puis   le r soudre, en g n ral par une m thode it rative plut t que directe   cause de sa taille. Plusieurs choix de fonctionnelle sont possibles. La m thode de collocation stipule simplement que l' galit  1.5 doit  tre v rifi e en un point x_i associ    chaque f_i . Plus g n ralement, on peut convoluer chaque membre de l' galit  par une s rie de fonctions de pond ration w_i . Vue sous cet angle, la collocation correspond   l'emploi de fonctions de Dirac. Une autre possibilit  est de prendre les fonctions de base f_i elles-m mes comme w_i , c'est la formulation de Galerkin [Zat92]. Associ e   l'utilisation de fonctions de base et de r flectances constantes sur les facettes, la m thode de Galerkin correspond   la formulation originale de la m thode de radiosit .

Dans cette formulation, le coefficient couplant deux facettes dans le syst me lin aire obtenu est le produit de la r flectance de l'une par un terme ne d pendant que de la g om trie : le facteur de forme. Cette particularit  est une cons quence directe de l'hypoth se lambertienne sur la r flexion de la lumi re. Le calcul de ces facteurs repr sente une grande partie du temps total de solution par la m thode de radiosit . Une fois qu'ils sont calcul s et stock s pour une g om trie donn e de la sc ne, les r flectances des objets peuvent  tre chang es et l' clairage correspondant recalcul    relativement peu de frais. A partir des valeurs des radiosits s sur

les surfaces d'une sc ne, on peut produire des images de celle-ci sans autres calculs que l' limination des parties cach es car pour les surfaces parfaitement diffuse, la radiosit  est proportionnelle   la luminance qui est la quantit  «mesur e» par l'oeil. Cette propri t  est s duisante mais place sur les algorithmes de maillage la contrainte suppl mentaire de ne pas introduire d'artefacts g nant pour l'oeil.

1.2 Algorithmes pour la radiosit 

Nous pr sentons maintenant les principaux algorithmes mis en  uvre dans le rendu d'une sc ne par la m thode de radiosit . Le calcul d'une image comporte plusieurs op rations qui sont souvent entrem l es en pratique :

- maillage des surfaces ;
- calcul des facteurs de forme ;
- r solution du syst me d' quilibre global ;
- production d'images.

Plusieurs de ces algorithmes serviront de base aux d veloppements des chapitres suivants et nous donnons ici les notations et explications n cessaires. Les techniques existantes de maillage adaptatif des surfaces seront quant   elles vues en d tail au chapitre suivant.

1.2.1 Calcul des facteurs de forme

Le facteur de forme pour deux facettes A_1 et A_2 s'obtient en int grnant le produit du facteur de forme diff rentiel par la fonction de visibilit  sur les deux surfaces des facettes. Il est donc donn  par l'int grale

$$F_{A_1, A_2} = \frac{1}{A_1} \int_{A_1} \int_{A_2} vis(dA_1, dA_2) \frac{\cos \theta_1 \cos \theta_2}{\pi r^2} dA_1 dA_2, \quad (1.6)$$

o  $vis(dA_1, dA_2)$ est la fonction de visibilit  entre dA_1 et dA_2 qui vaut 1 si les deux surfaces  l mentaires se voient et 0 sinon (voir figure 1.3). La solution exacte de 1.6 n'est possible que pour des configurations simples et sans obstruction. En pratique, on d compose d'abord l'int grale double en calculant l'int grale simple suivante

$$F_{dA_1, A_2} = dA_1 \int_{A_2} vis(dA_1, dA_2) \frac{\cos \theta_1 \cos \theta_2}{\pi r^2} dA_2. \quad (1.7)$$

pour plusieurs dA_1 sur la surface A_1 . Pour  valuer 1.7, on utilise des m thodes num riques d'int gration qui peuvent  tre class es en deux grands groupes :

- m thodes calculant directement l'int grale de surface ;

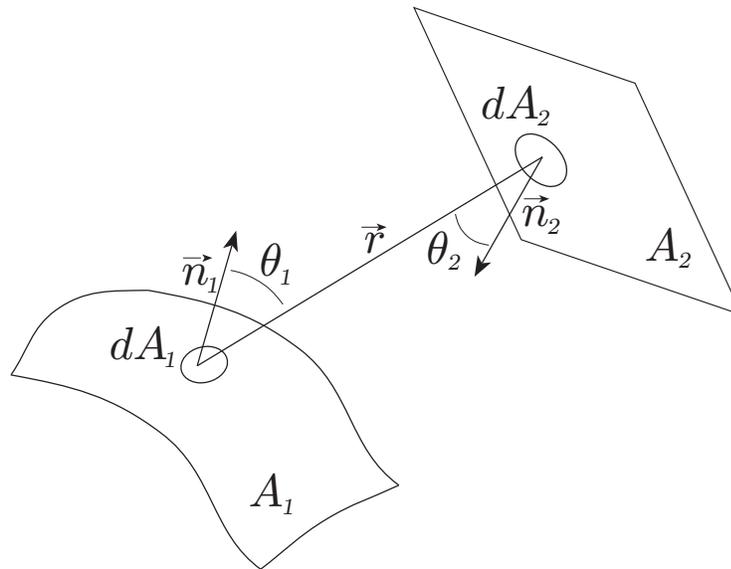


FIG. 1.3 - Géométrie pour les facteurs de forme.

- méthodes utilisant d'abord une transformation en intégrale de contour.

Intégrale de surface

Méthodes par projection Une première série de méthodes utilise l'équivalence entre le facteur de forme F_{dA_1, A_2} et l'aire de la projection orthogonale sur le plan tangent en dA_1 de la surface obtenue par projection radiale de la facette A_2 sur l'hémisphère unité centrée en dA_1 . Dans 1.7, le terme $\cos \theta_2 / \pi r^2$ correspond à la projection sur l'hémisphère et le terme $\cos \theta_1$ à la projection orthogonale sur le plan tangent. Cette relation, dite équivalent de Nusselt (voir figure 1.4), montre aussi que deux facettes ayant même projection sur la sphère ont même facteur de forme avec dA_1 .

En 1985, Cohen et Greenberg [CG85] utilisent comme surface intermédiaire de projection un demi cube (hémicube) placé au dessus de la facette pour déterminer quel objet est vu dans un ensemble de directions discrètes et donc quelle énergie y est incidente. La méthode de la radiosité est ainsi étendue à des environnements avec occlusion. En effet, les objets de la scène sont projetés sur les faces du cube en utilisant un algorithme de z-buffer pour prendre en compte la visibilité. Ces faces ont été au préalable subdivisées régulièrement en proxels (*projection elements*) dont on a précalculé les facteurs de forme. Le facteur d'une facette est alors obtenu en sommant les facteurs élémentaires des proxels qu'elle recouvre. Les implémentations de cet algorithme peuvent être rendues très efficaces par l'emploi d'un z-buffer cablé sur les stations de travail graphiques. D'autres méthodes à base de projection ont aussi été développées dont [SP89]. Comme indiqué plus haut, ces méthodes permettent d'évaluer l'intégrale intérieure F_{dA_1, A_2} de 1.6. L'intégrale

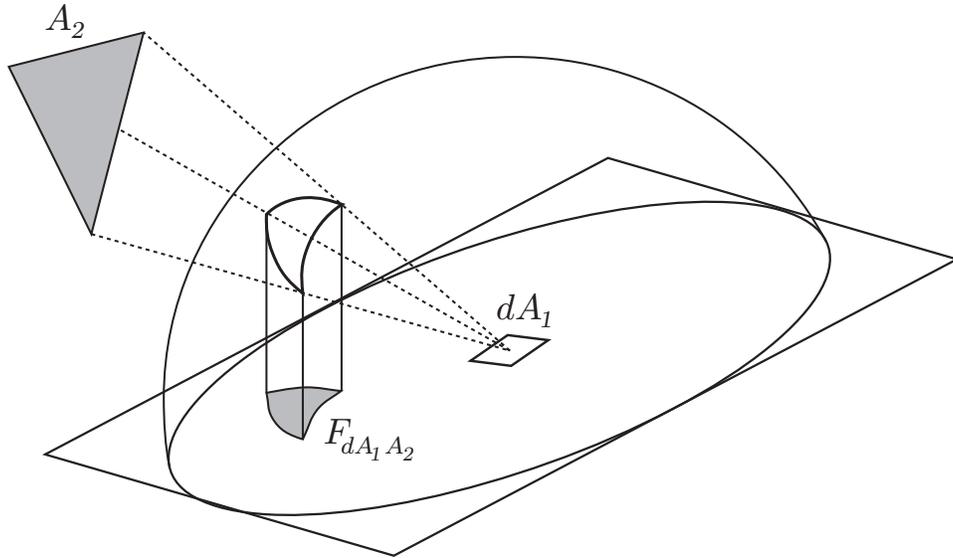


FIG. 1.4 - Équivalent de Nusselt pour le calcul des facteurs de forme.

extérieure est remplacée par une sommation finie. La facette A_1 est subdivisée en sous-facettes et un hémicube est calculé pour chacune. Cette approximation n'est valide que si l'intégrale intérieure varie peu dans la sous-facette associée à un hémicube donné ce qui implique que la taille de cette sous-facette est petite devant la distance aux autres objets.

Cette contrainte est l'un des deux principaux inconvénients des méthodes à base de projection car les objets les plus proches déterminent le nombre d'hémicubes à utiliser quelle que soit la géométrie du reste de la scène. D'autre part, la régularité de motif des proxels associée aux subdivisions régulières des surfaces réceptrices en facettes est source d'aliasage. Si certaines facettes sont trop petites, elles peuvent ne couvrir aucun proxel et créer des points d'ombre factices qui se répètent périodiquement sur les surfaces. Des solutions à ces problèmes ont été données dans [BRW89] mais elles peuvent devenir lourdes pour des scènes complexes ou pour des maillages adaptatifs fins.

Méthodes par lancer de rayons Une méthode de calcul différente a été proposée par Wallace *et al.* en 1989. Ici le point de vue est inversé et dA_1 est un petit élément de surface autour d'un point sur le *récepteur* alors qu' A_2 est la source. L'intégrale est approchée numériquement en divisant la source en plusieurs petits éléments de surface. Cependant, comme pour l'hémicube, cette approximation n'est valide que si l'aire de ces éléments est petite devant la distance entre dA_1 et A_2 . Pour contrôler la subdivision qui pourrait devenir extrêmement fine pour un récepteur proche de la source, on traite explicitement chaque sous-élément de A_2 comme ayant une aire finie. Pour cela, on utilise le facteur de forme entre

un disque et un  l ment de surface diff rentiel se faisant face dont la valeur est connue analytiquement. Comme dA_1 et A_2 ne sont pas forc ment l'un en face de l'autre, on utilise une approximation qui consiste simplement   multiplier le facteur par le produit des cosinus des angles entre les normales aux disques et le vecteur les joignant. Pour une division de A_2 en n sous- l ments  gaux on obtient la formule :

$$F_{dA_1, A_2} = \frac{dA_1}{n} \sum_{i=1}^n \frac{\cos \theta_{1_i} \cos \theta_{2_i}}{\pi r_i^2 + A_2/n} vis_i, \quad (1.8)$$

o  vis_i vaut 1 si l' l ment i est visible de dA_1 et 0 sinon. La formule du disque n'est qu'une approximation mais elle permet de r duire le nombre d' chantillons sur la source pour estimer rapidement le facteur de forme. Quand n augmente, F_{dA_1, A_2} tend vers la valeur exacte car A_2/n tend vers 0.

Les probl mes d'aliassage dont souffre l'h micube sont ici  limin s puisque le facteur de forme est calcul    chaque noeud du maillage. De plus, pour une pr cision de calcul donn e, le nombre d' chantillons utilis s est adapt  pour chaque point sur le r cepteur ind pendamment. Ces caract ristiques sont tr s int ressantes pour un algorithme de subdivision adaptative qui peut  tre amen    cr er des facettes de petite taille pour lesquelles les algorithmes utilisant une projection sont mal adapt s. Pour cette raison, les m thodes de calcul de d riv e et de maillage adaptatif pr sent es aux chapitres 3 et 4 sont fond es sur cette technique de calcul du facteur de forme.

Int grale de contour

L'int grale 1.7 peut aussi  tre transform e en int grale de contour. On commence par  crire

$$\begin{aligned} F_{dA_1, A_2} &= dA_1 \int_{A_2} vis(dA_1, dA_2) \frac{\cos \theta_1 \cos \theta_2}{\pi r^2} dA_2 \\ &= dA_1 \int_{A_2} \left(vis(dA_1, dA_2) \frac{\vec{n}_1 \cdot \vec{r}}{\pi r^4} \vec{r} \right) \vec{n}_2 dA_2 \end{aligned} \quad (1.9)$$

avec les notations de la figure 1.3. En utilisant le th or me de Stokes, on obtient ensuite l'int grale de contour

$$F_{dA_1, A_2} = dA_1 \int_{\partial_{vis} A_2} \frac{\vec{n}_1 \wedge \vec{r}}{2\pi r^2} d\vec{l}_2 \quad (1.10)$$

o  $\partial_{vis} A_2$ est le contour de la partie visible de A_2 et $d\vec{l}_2$ est un vecteur  l mentaire diff rentiel le long de $\partial_{vis} A_2$.

Comme l'int grale 1.7, cette int grale ne se calcule pas exactement dans le cas g n ral. Cependant, si le contour de la partie visible est une ligne polygonale

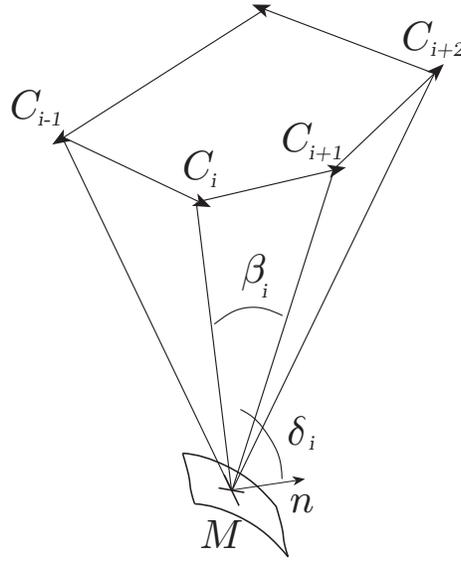


FIG. 1.5 - G om trie pour l'int gration de contour polygonal.

d finie par les points C_0, \dots, C_n (voir figure 1.5), $d\vec{l}_2$ reste constant par morceaux et l'int grale peut  tre mise sous la forme

$$F_{dA_1, A_2} = dA_1 \frac{1}{2\pi} \sum_{i=0}^{n-1} \beta_i \cos \delta_i, \quad (1.11)$$

o  β_i est l'angle (C_i, M, C_{i+1}) et δ_i est l'angle entre le plan d fini par M, C_i et C_{i+1} et la normale \vec{n} en M .

Pour les sc nes exclusivement compos es de poly dres, on peut donc calculer exactement le facteur de forme. Il faut pour cela d terminer le contour de la partie visible de la source. Plusieurs algorithmes utilisant l'int gration de contour ont  t  pr sent s qui diff rent essentiellement par la mani re de r soudre le probl me des parties cach es de la source. Dans [NN85], Nishita et Nakamae utilisent un algorithme d' limination des lignes cach es fond  sur la notion de visibilit  quantitative. [CF92] et [LTG92] utilisent eux des arbres BSP.

Au chapitre 6, nous pr sentons une m thode de calcul des facteurs de forme qui utilise l'int gration d'un contour polygonal approchant le contour r el et quelconque des objets. L'algorithme n'est donc plus limit  aux objets poly driques et peut traiter tous les objets accessibles au lancer de rayons qui est la primitive utilis e pour tester la visibilit .

1.2.2 R solution du syst me d' quilibre global

Apr s la phase de discr tisation, on obtient un syst me lin aire d' quations. Suivant la m thode employ e, les coefficients sont soit d riv s des facteurs de

forme entre facettes, soit le r sultat d'int grales plus complexes qui peuvent  tre calcul es par des formules de quadrature. Pour ces techniques, le lancer de rayon est bien adapt  aux tests de visibilit  point   point n cessaires. Les radiosit s constantes des facettes ou les coefficients de fonctions de bases plus compliqu es sont les inconnues de ce syst me. L'algorithme pr sent    l'origine par Goral *et al.* calcule tous les coefficients du syst me et le r sout par la m thode de Gauss. En 1988, Cohen *et al.* [CCWG88] proposent une autre organisation des calculs qui consiste    mettre successivement dans la sc ne l' nergie des  l ments de surface en poss dant le plus. Les deux  tapes de calcul des coefficients et de r solution sont alors m lang es et la production d'images de solutions partielles est aussi possible apr s chaque  mission d'une source. Des techniques de rajout de l' nergie non encore distribu e sous forme d'un terme ambiant ou d'utilisation simultan e de mod les plus simples d' clairage pour les lampes non encore trait es ont  t  d velopp es pour rendre les premi res images ainsi obtenues plus rapidement proches de la solution finale [CCWG88, PSV90].

1.2.3 Production d'images

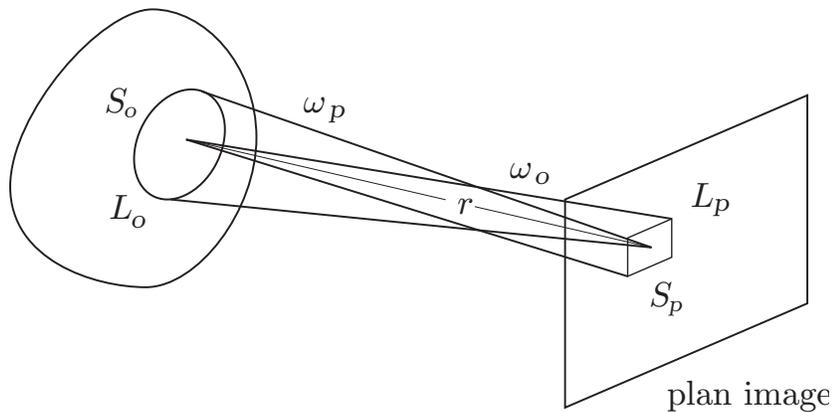


FIG. 1.6 - *Calcul de l' nergie incidente sur un pixel pour un mod le de prise de vue simple.*

La valeur   donner   chaque pixel de l'image doit  tre proportionnelle   l' nergie incidente sur sa surface. Cette  nergie est l'int grale sur le pixel de la luminance incidente et est donc proportionnelle   la luminance L_p sur le pixel. Or L_p est  gale   L_o , luminance sur l'objet vu par conservation du flux total dans le tube joignant S_p , surface du pixel et S_o , surface visible de l'objet (voir figure 1.6). Avec l'hypoth se lambertienne, la radiosit  est proportionnelle   la luminance et peut donc  tre utilis e directement pour l'affichage. La production d'une image   partir d'un maillage des surfaces repr sentant la fonction de radiosit  consiste donc simplement   d terminer les objets visibles dans le plan image. Tous les

algorithmes de parties cachées peuvent donc être utilisés et en particulier l'algorithme de z-buffer cablé. Pour des scènes suffisamment simples, on peut ainsi obtenir plusieurs images par seconde ce qui permet d'explorer virtuellement l'environnement en utilisant, par exemple, la souris pour changer le point de vue et donc se déplacer.

1.3 Vecteur de forme

La valeur de l'éclairement en un point d'une surface dépend de la position de la source, des obstacles éventuels et de la normale au point. Les variations qui doivent être capturées par le maillage des objets ont donc plusieurs origines. Habituellement, la quantité stockée aux sommets du maillage est la radiosité. Lors de la reconstruction, les valeurs de radiosité sont interpolées pour reconstruire une fonction continue sur les surfaces. Cette façon de procéder est correcte quand on a affaire à des surfaces planes, c'est-à-dire quand la normale est fixe. Mais pour les surfaces courbes, il est possible que la variation de l'éclairement soit due essentiellement aux changements de normale. Dans ce cas, un maillage fin est nécessaire alors que, placée au même endroit, une surface plane aurait besoin de beaucoup moins d'échantillons.

Pour découpler complètement éclairement et géométrie locale de la surface, nous proposons de stocker à la place de la radiosité un vecteur. Dans la formule du facteur de forme entre dA_1 et A_2 , la normale \vec{n}_1 peut être sortie de l'intégration pour obtenir

$$\begin{aligned} F_{dA_1, A_2} &= \left(dA_1 \int_{A_2} vis(dA_1, dA_2) \frac{\vec{r} \cdot \vec{n}_2}{\pi r^4} \vec{r} dA_2 \right) \cdot \vec{n}_1 \\ &= \vec{G}_{dA_1, A_2} \cdot \vec{n}_1. \end{aligned} \quad (1.12)$$

Nous proposons d'appeler \vec{G}_{dA_1, A_2} le *vecteur de forme* entre dA_1 et A_2 . Au lieu de stocker sur la surface une radiosité scalaire, on utilise un vecteur de radiosité obtenu en remplaçant le facteur de forme par le vecteur de forme. Au cours de la résolution du système d'équilibre, on est amené à accumuler les contributions des émetteurs traités successivement. L'addition en un même sommet des vecteurs de radiosité de plusieurs sources est bien sûr possible puisque le produit scalaire avec la normale est linéaire et les valeurs aux sommets du maillage seront finalement les mêmes que l'on utilise le vecteur ou bien le facteur de forme.

L'utilisation des vecteurs ne change donc pas les calculs de radiosité proprement dits mais permet une meilleure reconstruction des intensités sur les surfaces courbes en utilisant un maillage plus lâche. Le vecteur de radiosité en un point est obtenu par interpolation. On prend ensuite son produit scalaire avec la vraie normale au point. La différence avec l'interpolation habituelle est à rapprocher

de celle entre ombrage de Gouraud et ombrage de Phong. L'utilisation de la vraie normale à la surface en tout point permet, grâce à la variation continue de la normale à la reconstruction, d'éliminer les bandes de Mach qui apparaissent sinon près des maximums d'illumination.

Les méthodes de calcul du facteur de forme présentées plus haut permettent naturellement de calculer le vecteur de forme à la place du facteur de forme. La version «vecteur» de la formule du facteur de forme élémentaire 1.8 page 13 s'écrit

$$\vec{G}_{dA_1, A_2} = \frac{dA_1}{n} \sum_{i=1}^n vis_i \frac{\vec{r}_i \cdot \vec{n}_2}{\pi r_i^4 + A_2/n} \vec{r}_i, \quad (1.13)$$

et la formule d'intégration de contour polygonal 1.11 page 14 devient

$$\vec{G}_{dA_1, A_2} = dA_1 \frac{1}{2\pi} \sum_{i=0}^{n-1} \beta_i \vec{n}_i, \quad (1.14)$$

où \vec{n}_i est le vecteur normal au plan défini par M , C_i et C_{i+1} .

Dans les chapitres 3, 4 et 6, nous utiliserons partout le vecteur de forme à la place du facteur de forme. Ce choix est fait pour simplifier la présentation des algorithmes et toutes les formules où intervient le vecteur s'expriment facilement en terme de facteur. En pratique, seules les surfaces courbes sont vraiment concernées. Les surfaces planes doivent conserver le stockage scalaire plus compact et une implémentation plus simple et efficace. Dans le chapitre 2, nous examinons les différentes techniques de maillage qui ont été développées dans le cadre de la méthode de la radiosit .

Chapitre 2

Méthodes de maillage pour la radiosité

Comme nous l'avons vu au chapitre précédent, l'intensité lumineuse $B(x)$ sur les surfaces d'un environnement à l'équilibre est trop complexe pour être calculée et représentée exactement, même dans le cas restreint de surfaces parfaitement diffuses. En pratique, c'est donc une solution approchée du problème qui est recherchée. La dimension du problème est d'abord réduite par projection dans un espace de fonctions de dimension finie. On construit ensuite dans cet espace, une fonction $\tilde{B}(x)$, proche de la solution réelle. L'équation d'équilibre de l'énergie se met sous la forme d'un système d'équations linéaires couplant les coordonnées de $\tilde{B}(x)$ dans la base finie et dont les coefficients sont reliés aux facteurs de forme. La solution de ce système est la fonction approchée recherchée.

Pour construire le système d'équations et le résoudre, de nombreuses approximations sont faites : le calcul des facteurs de forme n'est jamais exact car les relations de visibilité entre objets rendent l'intégrale qui les définit trop complexe pour être calculée de manière symbolique, la solution du système est obtenue par des méthodes itératives, etc. Cependant, le choix de l'espace de fonctions dans lequel on travaille reste une étape décisive dans l'obtention d'un résultat à la fois correct et visuellement satisfaisant. En effet, l'illumination des surfaces purement diffuses, malgré l'absence de reflets et de caustiques, varie beaucoup à cause des ombres et des pénombres et présente de nombreuses discontinuités. Le nombre de fonctions de base pour une facette doit être suffisant pour capturer ces variations, faute de quoi même une résolution très précise du système ne donnera pas un résultat proche de la radiosité réelle.

L'indépendance de l'illumination par rapport au point de vue est un avantage de la méthode de radiosité qui permet de produire des images à un coût relativement faible une fois que la solution a été obtenue. On peut, par exemple, utiliser une station de travail graphique pour simuler un déplacement virtuel dans la scène en donnant à l'utilisateur le moyen de changer interactivement sa position et la direction de son regard. Mais cette indépendance signifie aussi que l'éclai-

rage des scènes est entièrement extrait de la représentation de $\tilde{B}(x)$ dans l'espace objet. Les défauts du maillage se retrouvent directement dans l'image et produisent des anomalies qui peuvent être très gênantes à l'oeil : fuites d'ombre ou de lumière, bandes de Mach, aliassage, etc. Ces points sont propres à la simulation de l'éclairage dans le but de produire des images et ne sont pas abordés dans le cadre des méthodes de type éléments finis. De nombreux algorithmes de maillage ont été développés depuis les débuts de la radiosité en image de synthèse. Nous présentons ici les principales approches.

2.1 Maillages polygonaux adaptatifs

Les premières méthodes proposées utilisent des maillages polygonaux adaptatifs. Les surfaces sont approchées par des polygones simples (triangles ou quadrilatères) : les facettes. Les fonctions de bases définies sur ces facettes sont constantes, linéaires ou bilinéaires et le raffinement du maillage se fait par subdivision des polygones en sous-polygones. Quand les radiosités utilisées sont constantes, on calcule la valeur pour un sommet en moyennant celles des facettes qui lui sont incidentes ce qui permet de reconstruire par interpolation une fonction continue pour l'affichage de la scène.

2.1.1 Maillage à plusieurs niveaux de détail

Dans les premières implémentations de la méthode de radiosité [GTGB84, CG85], c'est l'utilisateur qui doit, après obtention et visualisation d'une première solution, décider d'augmenter globalement la finesse du maillage et relancer les calculs pour améliorer la solution. La première méthode de maillage automatique et localement adaptative est présentée en 1986 par Cohen *et al.* [CGIB86]. Après calcul de la matrice de facteurs de forme et solution du système d'équations linéaires, les facettes sont parcourues et l'écart entre les radiosités de leurs sommets est comparé à un seuil pour décider de celles qui seront divisées plus finement.

Le point clef de l'algorithme est de faire jouer ensuite des rôles dissymétriques aux facettes de départ et aux sous-facettes obtenues par subdivision : c'est la notion de *substructuring*. Les nouveaux facteurs de forme sont calculés en plaçant des hémicubes sur les facettes uniquement mais en y projetant les sous-facettes pour obtenir des facteurs de forme de facettes à sous-facettes. On moyenne ensuite ces facteurs de forme pondérés par les aires des sous-facettes ce qui fournit des facteurs de facettes à facettes et un système linéaire de même dimension qu'au départ mais dont les coefficients sont plus précis. Ce système est résolu puis les radiosités des sous-facettes sont calculées par combinaison linéaire des facteurs de forme de facettes à sous-facettes avec les radiosités des facettes.

Ainsi pour N facettes et $M > N$ sous-facettes, on ne calcule et ne stocke que $M \times N$ facteurs de forme au lieu de M^2 et on résout un système de taille N au

lieu de M .

2.1.2 Maillage cohérent

En 1991, Baum *et al.* présentent un bilan assez complet des propriétés nécessaires d'un maillage pour les étapes successives d'un calcul de radiosité. Les propriétés identifiées sont :

- le modèle de la scène doit représenter un objet physique valide ;
- les intersections entre les faces doivent être explicites ;
- le modèle ne doit pas contenir de sommets en T , c.-à-d. de sommets situés strictement à l'intérieur d'une arête ;
- les facettes du maillage doivent être convexes et bien proportionnées ;
- la densité de sommets doit être suffisante pour détecter et représenter les variations d'illumination.

Plusieurs algorithmes permettant d'imposer et de maintenir ces propriétés sur un maillage régulier raffiné de manière récursive sont également proposés. Ils permettent le traitement automatique de modèles complexes et éliminent les problèmes de fuites d'ombre ou de lumière. Cependant les contraintes imposées créent des maillages d'une grande complexité, en particulier à cause de la nécessité d'utiliser des arbres quaternaires restreints. Le calcul explicite et l'utilisation des discontinuités de l'illumination, qui permettent de satisfaire toutes les propriétés énoncées plus haut, sont préférables pour les scènes polyédriques.

2.2 Maillages utilisant des arbres binaires de partition de l'espace

Dans [CF89], Chin et Feiner présentent un algorithme pour calculer les ombres projetées par une source ponctuelle dans un environnement polyédrique. Ils utilisent la structure d'arbre binaire de partition de l'espace ou arbre BSP, qui est une structure de données pour représenter une subdivision récursive de l'espace par des hyperplans. Les arbres BSP permettent entre autre :

- de parcourir efficacement les polygones d'une scène depuis une source ponctuelle dans un ordre vérifiant la propriété : si A est dans l'ordre avant B , B ne peut pas projeter d'ombre sur A ;
- de représenter un volume polyédrique dans une structure qui facilite le calcul d'intersection avec un polygone.

Leur méthode consiste à parcourir les polygones depuis la source pour calculer l'intersection de chacun avec le volume d'ombre préexistant et maintenu dans un arbre BSP, dit arbre SVBSP (*Shadow Volume BSP*). Le polygone est ainsi découpé en fragments éclairés et à l'ombre et la pyramide d'ombre de chaque fragment éclairé est ajoutée à l'arbre SVBSP pour le tenir à jour.

En 1990, Campbell et Fussell [CF90] utilisent cette méthode pour calculer le maillage d'une scène pour la radiosité. Les sources étendues sont subdivisées en éléments suffisamment petits pour être vus sous un angle solide inférieur à un seuil fixé depuis le polygone le plus proche. Chaque élément est ensuite utilisé comme une source ponctuelle de la méthode de Chin et Feiner, ce qui permet de subdiviser la scène en fragments qui voient chacun un nombre fixe d'éléments. Le problème de la visibilité étant ainsi réglé, le facteur de forme de chaque fragment est calculé en accumulant les contributions des éléments de la source qui lui sont visibles.

Cette méthode est la première à produire un maillage non régulier et tenant compte de la géométrie des obstacles. Cependant, elle possède de nombreux défauts :

- le maillage ne correspond pas aux vraies discontinuités de l'illumination ;
- la présence d'un objet proche de la source provoque une division fine de celle-ci et augmente la complexité du maillage de toute la scène ;
- l'utilisation de l'arbre BSP pour faire à la fois la subdivision et le calcul de la partie visible de la source donne des pénombres aliassées (voir [Cam91]).

Une version améliorée de l'algorithme qui utilise les discontinuités est présentée dans [Cam91]. La complexité du maillage est réduite mais les discontinuités ne sont vues que comme structure d'accélération et leur importance pour la reconstruction n'est pas soulignée.

2.3 Tracé des discontinuités

Sur une surface lisse, l'illumination due à une source étendue totalement visible est une fonction C^∞ . En présence d'obstacle, des discontinuités de la fonction et de ses dérivées apparaissent à cause des changements de surface visible de la source. Les premiers tracés de discontinuités sont présentés dans [NN83]. Dans [Hec91], Heckbert utilise des résultats sur les calculs de graphes d'aspect d'un environnement polyédrique [GM90] pour énumérer toutes les discontinuités réalisables dans ce cas. Celles-ci sont de deux types :

- EV (pour Edge-Vertex), définie par un sommet et une arête. La surface de discontinuité associée est un secteur angulaire plan qui intersecte les objets selon des segments de droites.

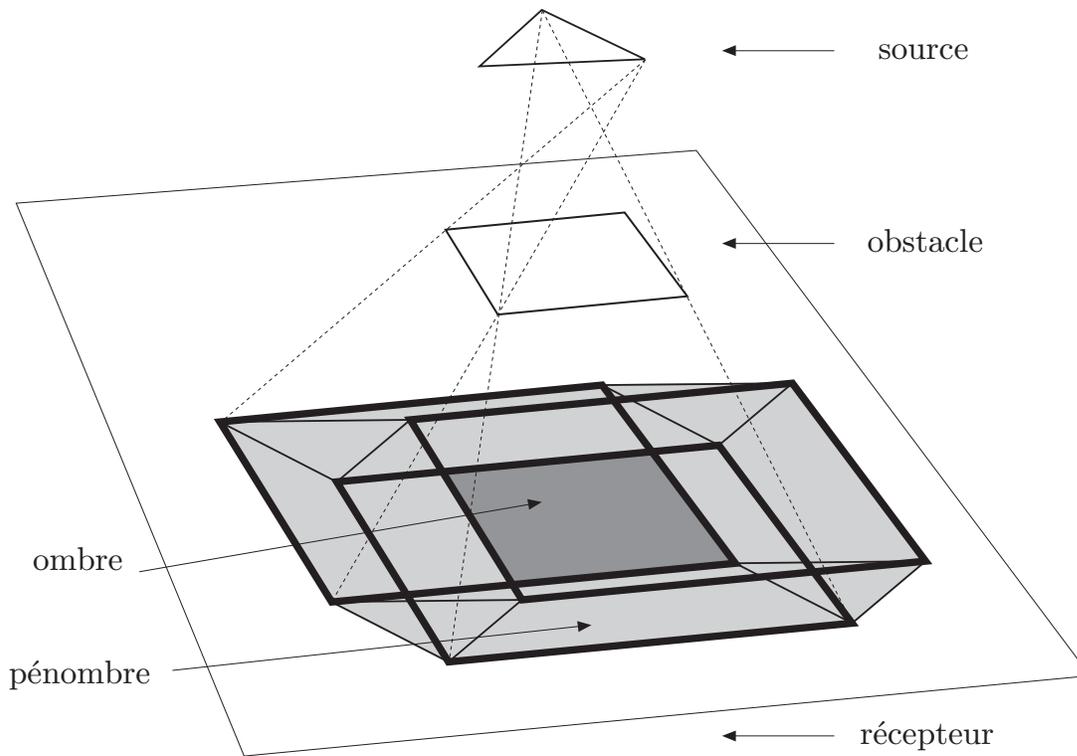


FIG. 2.1 - *Discontinuités pour une source triangulaire et un obstacle rectangulaire (adapté de [NN85]). Les discontinuités au tracé plus épais sont créées par un sommet de la source et une arête de l'obstacle, les autres par une arête de la source et un sommet de l'obstacle.*

- EEE (pour Edge-Edge-Edge), définie par trois arêtes en position générale. La surface de discontinuité associée est une quadrique qui intersecte les objets selon des segments de coniques.

La figure 2.1 montre les discontinuités pour une source triangulaire et un obstacle rectangulaire. Seuls des segments de type EV y sont visibles puisque la scène ne contient qu'un obstacle. La figure 2.2 montre un segment de discontinuité conique. Ces discontinuités sont en général dans la dérivée seconde mais la coïncidence de deux telles courbes (par exemple deux segments EV provenant d'une même arête et de deux sommets tous dans un même plan) peut engendrer une discontinuité dans la dérivée première. Comme indiqué dans [Hec91], des discontinuités des dérivées de degrés plus élevés sont obtenues quand la répartition d'énergie de la source est elle-même discontinue; c'est le cas pour les modèles d'illumination globale dans lesquels toutes les surfaces sont potentiellement émettrices. La notation D^i est introduite pour désigner une discontinuité de la i^e dérivée ou de la fonction elle-même pour D^0 . En pratique, seules les discontinuités de bas degré sont vraiment importantes à l'oeil et plusieurs algorithmes les introduisant explicitement dans le maillage ont été développés.

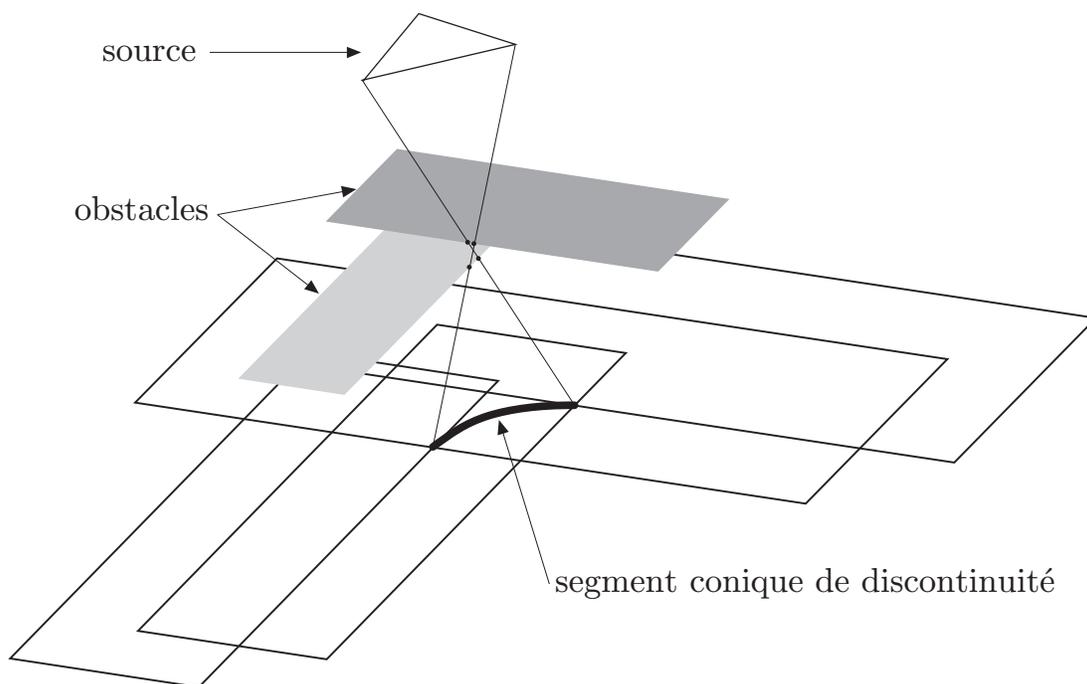


FIG. 2.2 - *Discontinuité de type EEE engendrée par une arête source et deux arêtes d'obstacles rectangulaires. Les autres discontinuités n'apparaissent pas toutes sur la figure.*

Les implémentations de Chin et Feiner [CF92] et de Lichinski *et al.* [LTG92] ignorent les événements de type EEE. Toutes les discontinuités retenues sont alors des segments de droites ce qui permet d'utiliser des arbres BSP pour coder la subdivision des polygones. Ce codage fonctionne mais crée beaucoup de polygones inutiles car la coupure ne se limite pas à la longueur de la discontinuité. Les fragments obtenus peuvent encore être subdivisés pour assurer une illumination correcte. Un traitement moins brutal passerait par la construction du vrai graphe des discontinuités mais se pose alors des problèmes de robustesse dus aux imprécisions numériques et à la nécessité d'obtenir un objet de topologie cohérente.

Le cas des surfaces courbes n'a encore jamais été abordé, même si les événements visuels possibles sont connus théoriquement [Rie87]. En fait, même si on se limite à des source polygonales convexes, les surfaces courbes posent de nombreux problèmes pratiques d'implémentation :

- les surfaces critiques des événements EV pour lesquels le sommet appartient à la source deviennent plus complexes. La surface critique est un cône s'appuyant sur le contour d'occlusion de l'objet vu depuis le sommet. Un calcul numérique approché de ces contours est possible mais coûteux [PK89] et il faut ensuite intersecter le cône avec d'autres objets pour trouver les

courbes de discontinuité ;

- le maillage obtenu est alors curviligne et l'imposition du bon degré de continuité au travers de frontières courbes est un problème d'interpolation difficile ;
- l'utilisation de polyèdres permet un algorithme plus efficace car les faces peuvent être partiellement ordonnées depuis un sommet de la source. Ce n'est pas le cas des surfaces courbes, qui elles peuvent contenir des parties intrinsèquement concaves.

Dans le chapitre 7, nous proposons une méthode qui utilise certaines des discontinuités créées par des surfaces courbes. Plutôt que de construire un maillage sur les objets, on calcule une approximation discrète des discontinuités dans l'espace image qui est ensuite utilisée pour accélérer le calcul d'ombrage d'un lancer de rayons.

2.4 Textures de radiosité

En 1990, Heckbert propose une structure de données complètement différente pour représenter l'illumination : les textures de radiosité [Hec90]. Sa méthode n'est pas limitée aux surfaces diffuses mais peut également traiter des surfaces dont la fonction de réflectance bidirectionnelle est une combinaison de réflexion et de transmission diffuses et purement spéculaires. Il n'y a pas non plus de limitation aux polyèdres et les surfaces courbes peuvent être traitées sans être polygonalisées.

Les textures de radiosités sont des arbres quaternaires dans l'espace des paramètres des surfaces. Dans chaque feuille sont accumulés le nombre de photons incidents et leur énergie totale. Ces textures ne sont associées qu'aux surfaces au moins partiellement diffuses. L'algorithme de calcul de l'illumination consiste à lancer des rayons depuis la source en déposant le photon associé et son énergie dans la première texture de radiosité rencontrée après d'éventuelles réflexions sur des surfaces spéculaires. Les rayons lancés depuis la source utilisent aussi une structure d'arbre quaternaire. Après l'envoi des rayons d'un niveau de l'arbre, les feuilles des textures ayant trop d'énergie sont subdivisées. Un critère est aussi utilisé pour décider de la subdivision de l'arbre des rayons. Puis tous les rayons sont à nouveau lancés dans les textures plus fines. Pour produire des images, on emploie un lancer de rayons sans rayon d'ombrage puisque l'illumination des surfaces diffuses est extraite des textures.

Cet algorithme permet en théorie de traiter des surfaces courbes avec une fonction de réflexion plus évoluée que la loi de Lambert et simule des transports de lumière mêlant réflexion diffuse et spéculaire. En pratique, cependant, de très nombreux rayons et des textures fines sont nécessaires pour réduire le bruit dans

les images à un niveau acceptable, ce qui rend l'algorithme très coûteux en mémoire et en temps de calcul.

2.5 Échantillons d'illumination en 3D

Ward et al. [WRC88] traitent le même type de surfaces qu'Heckbert mais avec une approche encore différente. Leur méthode utilise les techniques d'intégration de Monte Carlo pour calculer l'intensité de chaque pixel de l'écran. Un rayon est lancé dans la scène et si il rencontre un objet, on calcule, au point M d'intersection du rayon, l'éclairage direct des lampes et les contributions de la partie spéculaire et de la partie diffuse de la fonction de réflectance de la surface. C'est le dernier terme qui pose problème car il nécessite de tracer plusieurs dizaines de rayons pour intégrer l'intensité incidente dans l'hémisphère au dessus du point. Comme l'algorithme d'évaluation est récursif, le nombre de rayons à lancer explose.

La remarque faite par Ward et al. est que l'éclairage diffus des surfaces varie assez lentement à la surface des objets. Ainsi on peut se contenter de ne le calculer vraiment que pour quelques points bien choisis et utiliser ensuite ces valeurs pour le reconstruire ailleurs. L'algorithme utilise un octree d'échantillons dans l'espace. Ces échantillons sont situés sur les objets mais ne leur sont pas liés et la valeur stockée est l'intensité incidente et non réémise. Quand l'éclairage en un point doit être calculé, on cherche d'abord des échantillons proches dans l'octree. Le critère de proximité combine distance entre les points et angle entre les normales. Si suffisamment d'échantillons sont proches, on calcule une moyenne pondérée de l'illumination incidente qui est ensuite multipliée par la réflectance au point. Sinon le calcul complet est effectué et le nouveau point est ajouté dans l'octree. Au chapitre 4, nous présenterons une méthode de reconstruction semblable mais utilisée dans le cadre d'une radiosité progressive et avec des maillages liés aux objets.

2.6 Bases de fonctions globales

Cette méthode représente la radiosité sur une surface en la décomposant sur une base de fonctions polynômiales et a été proposée par Zatz en 1992 [Zat92]. La réduction de dimension du problème est obtenue ici en ne gardant qu'un sous-ensemble fini des fonctions de base, c'est-à-dire en ne conservant les fonctions que jusqu'à un degré maximal fixé. La manière de raffiner la représentation de la radiosité marque une différence importante avec les méthodes précédentes. Jusque là, la fonction était construite par morceaux et la solution était affinée en subdivisant la facette en sous-facettes pour obtenir des éléments de support plus petit. Ici, on rajoute des fonctions de bases de degré plus élevé mais le support de toutes ces fonctions est le même. Cette différence d'approche est

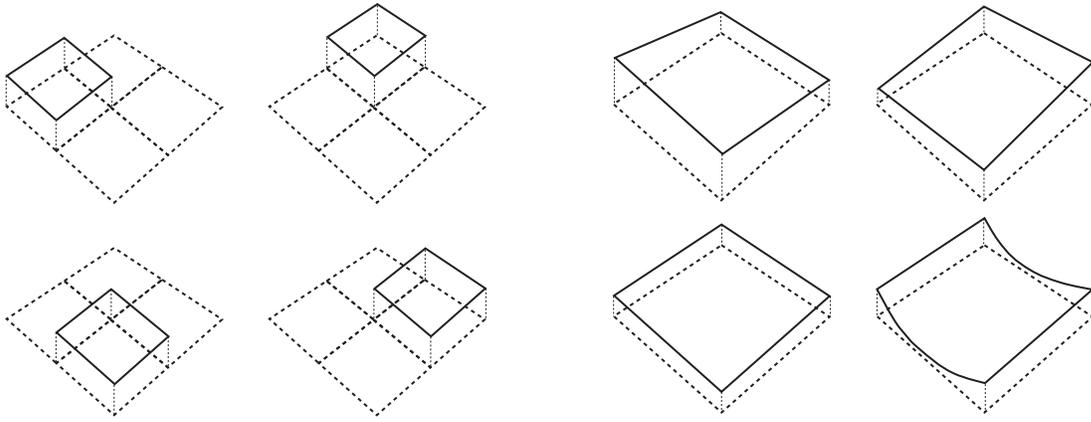


FIG. 2.3 - *Comparaison entre les méthodes de maillage habituelles pour lesquelles le raffinement se fait en divisant le support des fonctions (à gauche) et la méthode des fonctions globales pour laquelle on ajoute des fonctions de degré plus élevé (à droite).*

illustrée par la figure 2.3 qui est adaptée de [Zat92]. Ce choix permet de réduire la taille de mémoire nécessaire au codage de l'intensité. En effet, les fonctions de la base utilisée sont les mêmes pour toutes les facettes de l'environnement. N'étant pas liées à la géométrie des objets, elles ne requièrent pas de structures de données (arêtes ailées ou tri-quadtrees) pour coder la subdivision sur chacun. Chaque facette est simplement munie d'un vecteur représentant les coordonnées de sa fonction de radiosité dans la base. Il est à noter que les fonctions peuvent être définies dans l'espace des paramètres ce qui permet de traiter les surfaces courbes.

Les bases de fonctions utilisées sont des polynômes de Legendre et de Jacobi, ces derniers étant réservés aux facettes dont l'intensité présente une singularité due à une facette adjacente (coin). Les facteurs de forme sont remplacés par des facteurs de transfert entre fonctions de base sur les différentes facettes qui sont intégrés par quadrature. Un problème important de cette technique est le traitement des discontinuités des ombres et pénombres. Si elles ne sont pas introduites explicitement, elles provoquent des oscillations gênantes de l'intensité (phénomène de Gibbs) même pour un nombre relativement élevé de fonctions de base. Deux solutions sont proposées : le maillage préalable de la scène suivant les discontinuités qui à l'inconvénient de limiter la méthode à des environnements polyédriques et l'utilisation de masques d'ombre qui sont des textures régulières codant la fraction de source visible pour les surfaces où cela est nécessaire.

2.7 Méthodes hiérarchiques

Les algorithmes de radiosité hiérarchiques [HS90, HSA91, SAS92] ne sont pas à proprement parler des méthodes de maillage. Ils utilisent en fait des subdivisions récursives des facettes de type arbre quaternaire telles que celles décrites au §2.1.2. Cependant le résultat de ces algorithmes est intéressant. En effet, il montre que les problèmes de l'exactitude de la solution d'illumination globale d'une part et d'aspect des images de cette solution pour un observateur d'autre part placent des contraintes très différentes sur la méthode de résolution à employer.

Le principe des techniques hiérarchiques est d'optimiser les calculs d'illumination globale en essayant de consacrer à chaque transfert de lumière entre deux facettes un temps proportionnel à la quantité d'énergie transportée. Pour cela, les facettes sont subdivisées récursivement et le transport entre deux facettes se fait en suivant des liens entre les noeuds des arbres de subdivision respectifs. Le choix des noeuds à relier se fait en fonction de la quantité estimée d'énergie y transitant et d'une notion d'importance dans le cas de [SAS92]. Un lien trop grossier peut être raffiné en descendant dans l'un des arbres de subdivision. Ces méthodes sont inspirées de celles développées pour résoudre de manière approchée et efficacement le problème à n corps pour un grand nombre d'objets. Hanrahan et Salzman montrent dans [HS90] que le nombre de liens entre deux facettes est en $O(n)$ pour une division en n sous-facettes alors que la solution naïve est en $O(n^2)$. La solution hiérarchique est donc nettement plus efficace pour des facettes de départ assez grandes qui doivent être raffinées. Le problème de la création de liens entre groupes de petites facettes n'est pas encore complètement résolu mais une première approche est présentée par Rushmeier *et al.* dans [RPV93].

Les images obtenues par les techniques décrites précédemment ne sont pas satisfaisantes car les bords des ombres présentent de nombreux artefacts. Il faut donc ajouter des règles supplémentaires pour forcer le raffinement des liens et donc du maillage dans les pénombres [HSA91]. En fait, les critères employés par les méthodes hiérarchiques ne tiennent pas compte :

- de l'importance de la forme des ombres dans l'image ;
- de la sensibilité de l'oeil au contraste plutôt qu'à la valeur absolue de l'illumination.

L'opposition est nette avec les méthodes adaptatives de maillage qui raffinent aux frontières de l'ombre et de la pénombre sans tenir compte des quantités d'énergie mises en jeu.

2.8 Bilan

Nous pouvons tirer plusieurs conclusions à partir des inconvénients et avantages des méthodes de maillage existantes. Pour les scènes polyédriques, l'utilisa-

tion des discontinuités parait être une solution préférable aux maillages réguliers adaptatifs, aussi bien en ce qui concerne le temps de calcul que l'occupation mémoire. La méthode a cependant quelques inconvénients :

- le nombre de facettes créées ne peut pas être fixé facilement alors qu'une profondeur maximum de subdivision permet de contrôler indirectement ce nombre dans une méthode de raffinement récursif ;
- l'implémentation robuste du vrai maillage de discontinuités comprenant les événements EEE est difficile ;
- les maillages dûs à des sources différentes n'ont *a priori* pas de facettes communes et leur union augmente le nombre total de facettes ;

A cause de la dernière remarque, la méthode de maillage par discontinuités ne peut pas être appliquée à toutes les étapes d'une solution progressive. Nous verrons au chapitre 5 que la composante indirecte de l'éclairage varie assez lentement et que sa précision dépend peu de la finesse de représentation de l'éclairage direct. La solution est donc de calculer la composante indirecte sur un maillage lâche indépendamment de la composante directe due aux sources primaires. Un tel calcul peut d'ailleurs être effectué par une méthode hiérarchique simple, c'est-à-dire sans raffinement des ombres et pénombres.

Pour les scènes comportant des surfaces courbes, le problème du calcul des discontinuités est plus complexe et rend l'utilisation d'un maillage adaptatif nécessaire. Pour éliminer tous les artefacts provenant du maillage dans les ombres, on peut recalculer celles-ci au moment du calcul de l'image en ne retenant du maillage que la partie indirecte de l'illumination et en lançant des rayons d'ombrage vers les sources primaires. Dans les deux derniers chapitres, nous verrons deux méthodes d'accélération de ce calcul.

Dans le chapitre 3, nous commençons par proposer une méthode de calcul des dérivées premières de l'illumination fondée sur la méthode d'estimation du facteur de forme par lancer de rayons [WEH89] et qui ne requiert pas de calculs de visibilité supplémentaires. Ces dérivées seront ensuite utilisées dans le chapitre 4 qui présente une nouvelle méthode de maillage capable de traiter les surfaces courbes sans les approcher d'abord par des polygones.

Chapitre 3

Dérivées premières de l'illumination

Les méthodes de calcul de facteur de forme à base de projection (hémicube) fournissent des facteurs de forme entre facettes de la scène. Pour la reconstruction, des valeurs aux sommets des facettes sont obtenues en moyennant les contributions des facettes incidentes. Par contre, les méthodes à base de lancer de rayon fournissent une valeur ponctuelle du facteur de forme c'est-à-dire le facteur de forme entre la source et un petit élément de surface autour d'un point du récepteur. Ces calculs nécessitent pour les sources importantes l'envoi de nombreux rayons et l'information fournie par ces rayons dans la pénombre n'est pas entièrement utilisée par le simple calcul du facteur de forme. Dans la pénombre, la forme de la partie cachée de la source et la distance de l'obstacle doivent permettre d'estimer les variations du facteur de forme autour du point. Nous verrons au chapitre 4 que la connaissance de la dérivée apporte une amélioration du critère de subdivision adaptative et une meilleure reconstruction de la fonction sur la surface.

3.1 Utilisation du vecteur de forme

Dans [WH92], des gradients sont calculés pour aider à la reconstruction de l'éclairage indirect. Avec cette méthode et contrairement au cas qui nous intéresse ici, un échantillon d'illumination placé sur un objet est susceptible d'être utilisé pour reconstruire l'intensité en un point situé sur un autre objet. On ne peut donc plus définir de système commun de coordonnées locales (u, v) pour l'échantillon et le point où on reconstruit la valeur. On n'a pas affaire à une fonction de deux mais de cinq variables : trois variables pour la position dans l'espace et deux pour la direction de la normale. Ceci explique l'emploi par Ward et Heckbert de deux gradients séparés, un gradient de translation pour les variations des coordonnées d'espace et un gradient de rotation pour les variations de la normale.

Pour la reconstruction sur une même surface, il n'est pas nécessaire de séparer les variables car on peut voir le facteur de forme avec une source comme une fonction des coordonnées locales (u, v) . Cependant, comme annoncé au §1.3, on préfère utiliser le vecteur de forme qui permet à la subdivision et à la reconstruction de s'affranchir en partie des problèmes liés aux variations de la normale. De plus, on calcule les dérivées par rapport aux coordonnées globales (x, y, z) plutôt que par rapport aux coordonnées locales (u, v) . Ce choix permet en effet de reconstruire l'intensité continûment aux travers de surfaces lisses construites par recollement. Il suffit pour cela d'une notion de voisinage entre morceaux mais on n'a pas besoin de définir de coordonnées locales cohérentes sur chaque demi surface. Cette construction peut d'ailleurs être impossible si on utilise des patches généralisés à n côtés [LD90] ou des surface à continuité G_1 mais non C_1 .

3.2 Origine des variations d'illumination

La différence de vecteur de forme entre deux points M et M' peut se mettre sous la forme

$$\int \int_{A_j^{M'/M}} \frac{\vec{n}_j \cdot \vec{r}'}{r'^4} \vec{r}' dA_j + \int \int_{A_j^M} \left(\frac{\vec{n}_j \cdot \vec{r}'}{r'^4} \vec{r}' - \frac{\vec{n}_j \cdot \vec{r}}{r^4} \vec{r} \right) dA_j. \quad (3.1)$$

où A_j^M est la partie de A_j visible depuis M et $A_j^{M'/M}$ est la partie de A_j visible depuis M' mais pas depuis M . Cette décomposition illustre les deux causes de la variation d'illumination dans la pénombre. Quand on se déplace sur le récepteur de M vers M' (voir figure 3.1), la partie visible de la source change (premier terme) ainsi que l'angle solide sous lequel elle est vue (deuxième terme).

Une méthode naturelle pour estimer les dérivées d'une fonction f dont on n'a pas d'expression symbolique mais dont on sait calculer les valeurs point par point est l'utilisation de différences finies. On approche, par exemple, la dérivée première en utilisant la différence entre les valeurs pour deux points voisins :

$$f'(x) \simeq \frac{f(x+h) - f(x-h)}{2h}. \quad (3.2)$$

Dans le cas présent, il faut tenir compte du fait que la valeur calculée de la fonction n'est pas exacte mais quantifiée par le nombre d'échantillons et donc de rayons utilisés pour estimer la partie visible de la source. Dans la pénombre, la fonction réellement obtenue présente donc des marches d'escalier (voir figure 3.7, page 40). Pour obtenir une estimation correcte, il faut choisir une valeur de h et un nombre de rayons tels que la différence de valeur entre les deux points ait une signification. Ce choix dépend en fait de la position relative de l'obstacle entre la source et le récepteur et rend la méthode difficile à mettre en œuvre. Le coût supplémentaire en temps de calcul est également important.

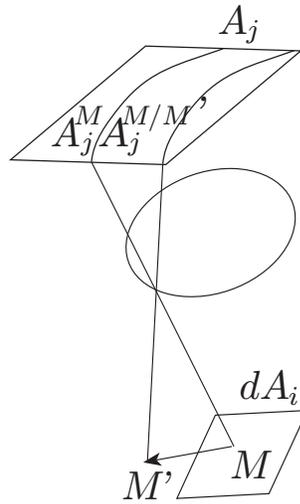


FIG. 3.1 - Géométrie de A_j^M et $A_j^{M'/M}$

Pour obtenir deux dérivées directionnelles, il faut lancer des rayons depuis trois points différents alors que c'est l'opération qui représente l'essentiel des calculs même pour des scènes ne comportant que peu d'objets. La figure 3.8, page 40, présente la dérivée obtenue par différence finie pour un choix de h donné. Un choix différent change la fréquence d'aliassage mais ne permet pas de l'éliminer partout. Nous proposons une utilisation différente des informations déjà présentes pour approcher les dérivées partielles du vecteur de forme sans lancer de rayons supplémentaires.

3.3 Calcul de la dérivée

En revenant à la décomposition présentée plus haut, nous allons d'abord voir que l'estimation du changement d'angle solide est facile. En effet, quelle que soit la formule utilisée (somme de facteurs élémentaires ou intégration de contour), l'approximation obtenue est assez bonne pour que sa dérivée soit elle-même proche de la dérivée de la vraie fonction. Pour estimer le changement de partie visible de la source, nous présentons ensuite une méthode qui calcule le déplacement sur la source de la frontière visible/cachée sans lancer de rayons supplémentaires.

3.3.1 Variation d'angle solide

Nous traitons ici le cas du facteur de forme élémentaire, celui de l'intégration de contour étant vu au chapitre 6. De la formule 1.8 page 13, on obtient le vecteur

de forme élémentaire entre dA_1 et dA_2

$$\vec{G}_{dA_1, dA_2} = \frac{\vec{r} \cdot \vec{n}_2}{\pi r^4 + dA_2} \vec{r}, \quad (3.3)$$

avec les notations de la figure 1.3 page 11. On obtient par dérivation partielle, par exemple par rapport à la coordonnée x

$$\partial_x \vec{G}_{dA_1, dA_2} = dA_1 dA_2 \frac{r^2(\pi r^2 + a)(\vec{n}_2 \cdot \vec{i}) - 2(2\pi r^2 + a)(\vec{r} \cdot \vec{i})(\vec{n}_2 \cdot \vec{r})}{r^4(\pi r^2 + a)^2} \vec{r}. \quad (3.4)$$

Comme pour le vecteur de forme, la contribution de chaque échantillon visible sur la source doit être accumulée pour obtenir la valeur totale. Cette expression est moins coûteuse à calculer qu'il ne paraît car des sous-expressions ont déjà été évaluées et le produit scalaire par \vec{i} consiste simplement à extraire la première coordonnée.

3.3.2 Variation de la partie visible

On utilise les notations de la figure 3.2 en considérant pour l'instant une source rectangulaire. Les vecteurs \vec{u} et \vec{v} forment une base orthonormée du plan de la source complétée par \vec{w} en base orthonormée de l'espace.

Dans le calcul du vecteur de forme, on a obtenu un arbre quaternaire d'échantillons sur la source. Un rayon a été lancé vers chacun de ces échantillons et éventuellement intercepté par un obstacle. Pour simplifier l'algorithme, on suppose que la source est subdivisée en une grille régulière de cellules rectangulaires. En pratique, on s'est contenté avant le calcul des dérivées de rendre l'arbre quaternaire complet en subdivisant les feuilles situées trop haut dans l'arbre. Cette étape est peu coûteuse car elle consiste simplement à recopier l'information d'une facette dans ses sous-facettes. Chaque cellule représente un facteur élémentaire et peut être vue ou non depuis le récepteur. Si elle est cachée, on dispose en plus de la distance de l'obstacle qui est la distance à laquelle le rayon associé a rencontré l'obstacle.

On parcourt ensuite les cellules par colonnes puis par lignes. Chaque couple (C_a, C_b) de cellules voisines suivant une direction se trouve dans l'une des trois configurations suivantes :

1. Les deux cellules sont vues.
2. Les deux cellules sont cachées.
3. Une cellule est vue et l'autre est cachée.

Dans le dernier cas, une frontière est détectée et on va calculer sa contribution aux différentes dérivées partielles (voir figure 3.2). Une première méthode de calcul

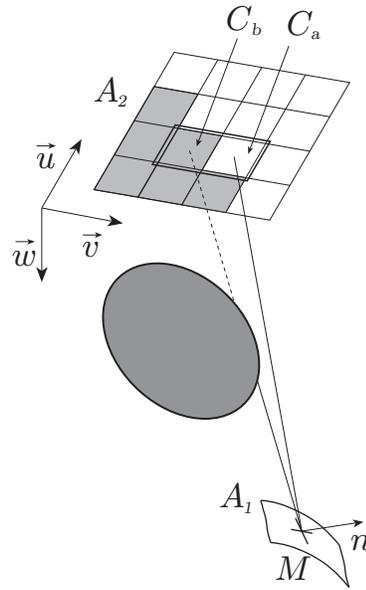


FIG. 3.2 - Couple de cellules (C_a, C_b) pour le parcours suivant les lignes

du gradient a été décrite dans [Ved92]. Nous présentons ici une méthode plus simple qui calcule les dérivées partielles du vecteur de forme. Pour cela, on utilise l'égalité

$$\partial_x \vec{G} = \Delta_{\vec{x}} \vec{G} = \lim_{h \rightarrow 0} \frac{\vec{G}(M + h\vec{x}) - \vec{G}(M)}{h} \quad (3.5)$$

qui relie dérivée partielle par rapport à une coordonnée et dérivée dans la direction du vecteur associé. Le bord de l'obstacle réel pouvant être compliqué, on va utiliser un obstacle proche et suffisamment simple pour que le calcul de sa contribution soit possible.

On suppose, par exemple, que C_a et C_b sont voisines suivant la direction \vec{v} , que C_a est vue alors que C_b est cachée et qu'on calcule la dérivée dans la direction de \vec{v} . Le principe est de calculer la différence de couverture de la source pour un petit déplacement de M à $M + h\vec{v}$. Ce calcul est simplifié si le bord de l'obstacle projeté sur la source est une droite perpendiculaire à \vec{v} qui se déplace en translation. On est alors amené à choisir comme obstacle un demi-plan dont le bord ne rencontre pas le plan de la source et est en fait parallèle à \vec{u} . De plus, ce demi plan doit se situer à peu près à la même distance que l'obstacle réel. On construit donc un obstacle virtuel représenté sur la figure 3.3 et défini par le vecteur \vec{u} et par le point O situé sur la droite MI à la distance d'interception du rayon associé à C_b .

Ainsi en M , O se projette sur I et l'obstacle sépare exactement C_a et C_b . Le vecteur de forme total pour les deux cellules est celui de C_a seule. Quand on se

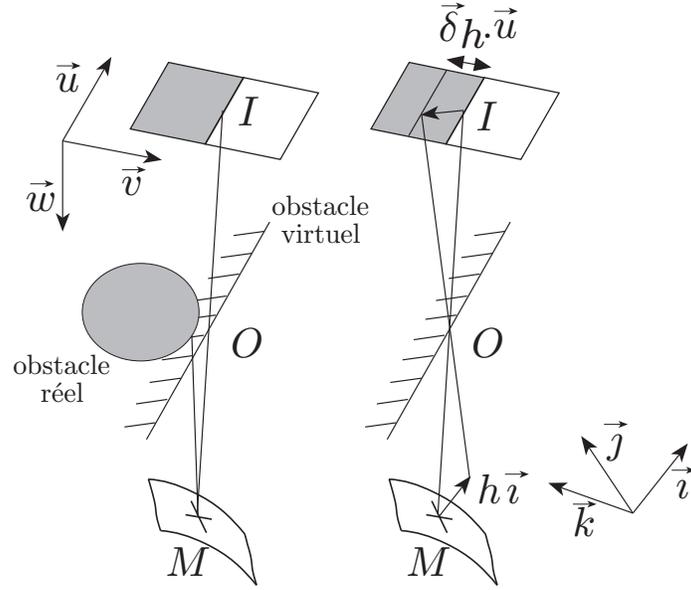


FIG. 3.3 - Géométrie de l'obstacle virtuel.

déplace de $h\vec{i}$, la translation δ_h de la projection de O sur la source est définie par :

$$\vec{OI} + \delta_h = \lambda(\vec{OM} + h\vec{i}) \quad (3.6)$$

$$\delta_h \cdot \vec{w} = 0. \quad (3.7)$$

Cette translation provoque un changement de visibilité pour une bande de l'une des deux cellules. La longueur de cette bande permet de calculer la variation de vecteur de forme comme une fraction de \vec{G}_a ou \vec{G}_b .

En faisant tendre h vers 0, on obtient

$$\lim_{h \rightarrow 0} \frac{\delta_h \cdot \vec{U}}{h} = \vec{IO} \cdot \vec{u} + \frac{\vec{OI} \cdot \vec{w}}{\vec{OM} \cdot \vec{w}} \vec{i} \cdot \vec{u}, \quad (3.8)$$

dont on déduit la contribution à la dérivée. Ces contributions sont accumulées pour chaque couple de cellules. La méthode qui a été présentée pour une source rectangulaire peut être étendue à toutes les sources polygonales, moyennant une triangulation de celles-ci. L'opération effectuée sur un couple de cellules pour estimer le changement de partie visible peut s'étendre à tout couple de facettes partageant une arête commune et il suffit donc de parcourir les arêtes de la triangulation pour accumuler leurs contributions éventuelles à la dérivée.

3.4 Résultats

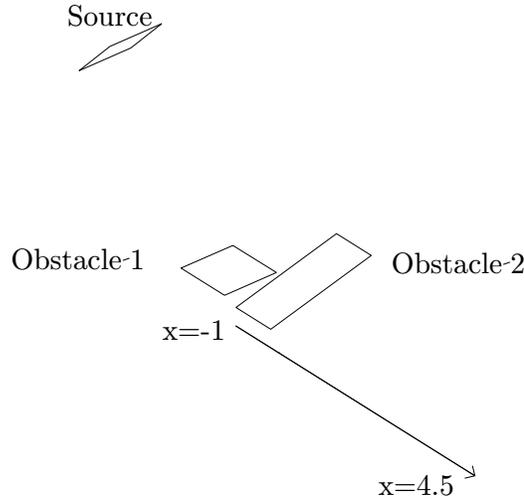


FIG. 3.4 - Géométrie pour les courbes.

Pour valider l'algorithme, on a utilisé la scène polygonale schématisée à la figure 3.4 et composée d'une source carrée et de deux obstacles rectangulaires. Les figures 3.5 à 3.9 montrent, pour différentes méthodes, les profils de facteur de forme et de dérivée partielle par rapport à x le long d'un segment tracé sur un plan récepteur et matérialisé par une flèche sur la figure 3.4. Les courbes de référence ont été obtenues par intégration de contour, technique qui grâce à sa grande précision permet de calculer les dérivées partielles par différences finies.

Sur les courbes 3.5 et 3.6, les discontinuités d'illumination ont été étiquetées A , B , C , D et E . Les intervalles $[A, B]$ et $[C, D]$ correspondent à la pénombre du premier obstacle et $[B, C]$ à son ombre. E marque le début de la pénombre du second obstacle. Comme le premier obstacle possède deux bords parallèles à ceux de la source, A , B , C et D sont des discontinuités de dérivée première, désignée par D^1 dans [Hec91]. E est une discontinuité de dérivée seconde, D^2 , car le second obstacle est incliné.

Les facteurs de forme de la courbe 3.7 ont été calculés en utilisant la méthode de Wallace avec 64 rayons par point. On note, dans la pénombre du premier

objet, les huit marches qui correspondent au changement de visibilité des huit rangées d'échantillons. La pente des marches correspond au changement d'angle solide alors que la partie visible reste constante. Les dérivées de la figure 3.8 ont été obtenues par différences finies à partir de cette courbe. Comme souligné au §3.2, la quantification de la courbe de départ entraîne un aliassage important de la dérivée. Le résultat peut être amélioré en changeant l'écart entre les points utilisés mais la valeur optimale dépend de l'emplacement de l'obstacle et l'aliassage n'est jamais totalement éliminé.

La méthode décrite à la §3.3 a été utilisée pour la figure 3.9. On retrouve dans les pénombres les marches de la courbe 3.7 mais la valeur de la dérivée reste proche de la référence. La courbe 3.10 montre l'erreur relative. Les pics de cette courbe au niveau des discontinuités A , B , C , D et E sont inhérents au placement des échantillons strictement à l'intérieur de la source. La largeur de la pénombre est ainsi légèrement réduite par rapport à la réalité créant une forte erreur relative aux transitions. En pratique, ce problème n'est pas crucial car l'erreur correspond à un simple décalage spatial. La dérivée est fautive mais il existe un point proche sur le récepteur possédant cette valeur. Ce point est d'autant plus proche que la pénombre est étroite et le décalage ne pose donc pas de problème dans le cas d'un maillage de radiosité pour lequel les sommets sont plus espacés.

Dans le chapitre suivant, nous présentons une méthode de maillage qui traite les surfaces courbes sans les remplacer par une collection de polygones et qui utilise les dérivées pour améliorer la reconstruction.

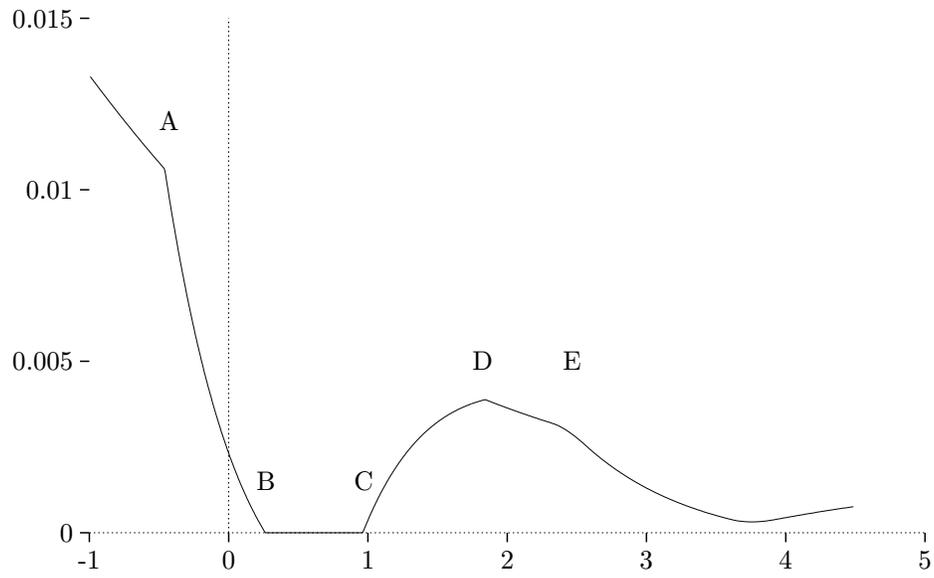


FIG. 3.5 - Courbe de référence du facteur de forme calculée par intégration de contour.

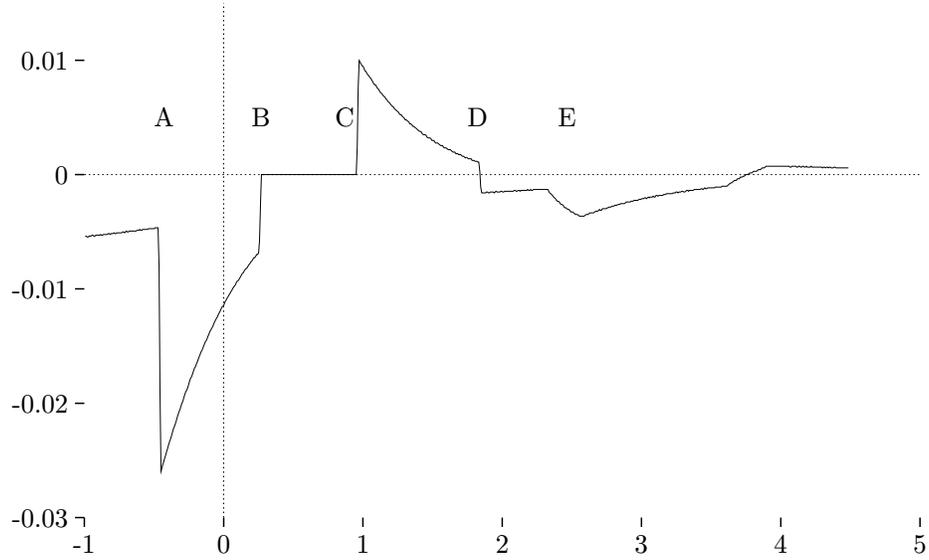


FIG. 3.6 - Courbe de référence de la dérivée partielle par rapport à x du facteur de forme calculée par intégration de contour.

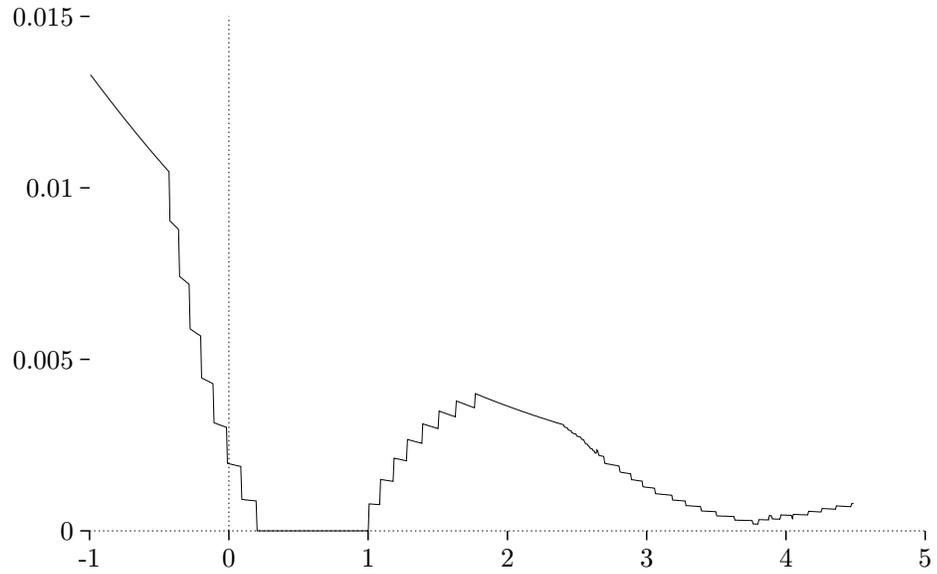


FIG. 3.7 - Courbe du facteur de forme calculée par la méthode de Wallace avec 64 échantillons par point.

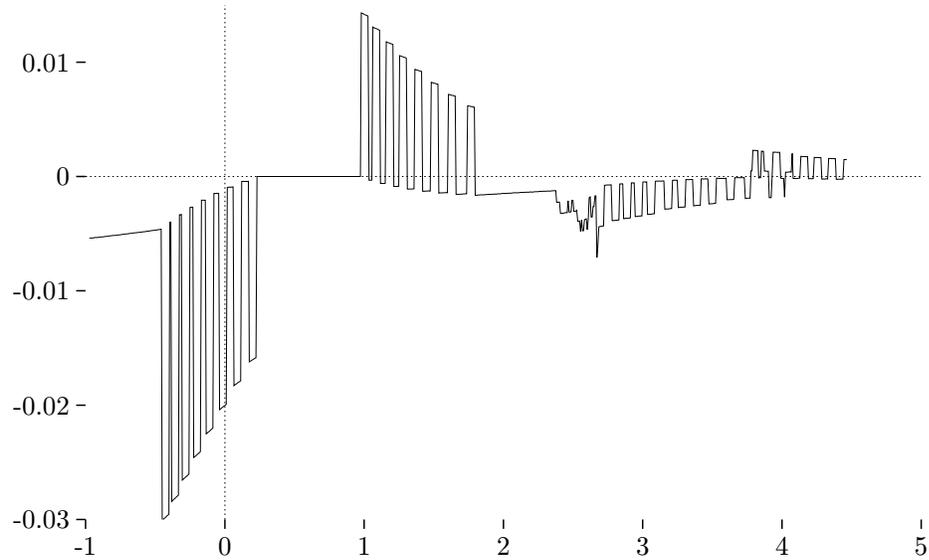


FIG. 3.8 - Courbe de la dérivée partielle par rapport à x du facteur de forme calculée par différences finies.

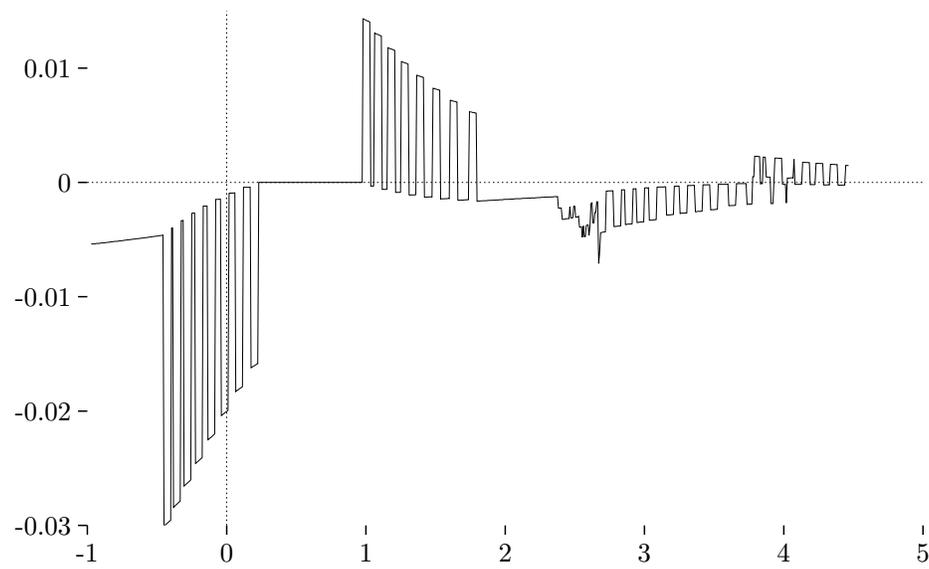


FIG. 3.9 - Courbe de la dérivée partielle par rapport à x du facteur de forme calculée par la méthode proposée.

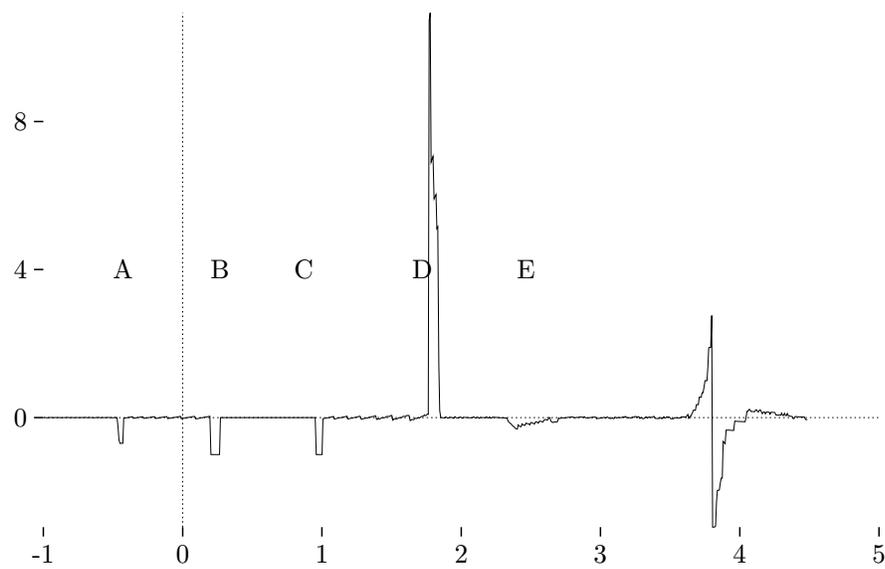


FIG. 3.10 - Erreur relative entre la méthode proposée et la référence.

Chapitre 4

Maillage adaptatif sans polygones

Dans un environnement exclusivement composé de polyèdres, le calcul d'un maillage de discontinuités permet d'obtenir une représentation très précise de l'illumination directe par les sources. En comparaison, un maillage adaptatif, plus simple à implémenter mais plus coûteux en temps de calcul et en espace mémoire, n'est rentable que si on accepte une représentation plus grossière des ombres. Cependant, l'introduction de surfaces courbes dans l'environnement change les données du problème. En effet, les courbes de discontinuités sur les objets sont alors de degré plus élevé et leur calcul exact n'est pas toujours possible. D'autre part, on ne peut pas utiliser une approximation polyédrique intermédiaire des objets sans faire exploser la taille du problème. Pour une source donnée, le nombre d'événements EV augmente comme $O(n)$ et celui d'événements EEE comme $O(n^2)$ où n est le nombre d'arête dans la scène or l'approximation précise d'une silhouette courbe requiert de nombreux polygones donc de nombreuses petites arêtes.

Dans cette situation, une technique de maillage adaptatif reste donc une solution intéressante. La plupart des méthodes de maillage existantes commence par approcher les surfaces courbes par un réseau polygonal. Deux exceptions notables sont l'utilisation des textures de radiosité par Heckbert [Hec90] et l'utilisation de bases orthogonales de polynômes définis sur l'espace des paramètres de la surface par Zatz [Zat92]. Nous avons vu au chapitre 2 les inconvénients de ces deux approches. Les textures de radiosité imposent une discrétisation très fine donc coûteuse pour obtenir une représentation fidèle et non bruitée de l'éclairage. Les bases globales obligent de recourir à un traitement séparé des ombres par des textures d'ombrage (*shadow masks*).

Dans ce chapitre, nous présentons une méthode de maillage adaptatif qui traite les surfaces courbes sans les rendre polygonales. Elle est utilisée par un algorithme de radiosité progressive [CCWG88] dans lequel on sélectionne successivement les objets ayant le plus d'énergie restant à émettre pour la redistribuer dans l'environnement. Dans ce cadre, les opérations à réaliser sur le maillage pour calculer une solution et l'afficher sont les suivantes :

- réception et stockage de la lumière venant des autres surfaces de l'environ-

nement ;

- émission de la lumière vers les autres surfaces de l’environnement ;
- reconstruction de l’éclairage en tout point de la surface.

Les méthodes de calcul de facteur de forme utilisant des projections rendent nécessaire la polygonalisation des surfaces. La contribution reçue par chaque échantillon dépend en effet du nombre de proxels que sa facette recouvre sur l’hémicube par exemple. L’emploi du lancer de rayons laisse lui plus de libertés quant au placement des échantillons. La technique de maillage décrite dans ce chapitre n’utilise pas de division des surfaces en facettes dans le but d’éviter :

- d’avoir à gérer une subdivision récursive avec partage de sommets ;
- de limiter le choix des fonctions de base à cause de la géométrie des facettes qui les supportent.

La solution présentée dans la suite utilise le voisinage spatial entre échantillons pour détecter les ombres et y raffiner le maillage. Par contre, les échantillons ne sont pas partagés par plusieurs facettes ce qui laisse plus de latitude pour les déplacer et obtenir un maillage plus régulier (§4.1.1). L’implémentation est également simplifiée et nous verrons que la contrainte de restriction de l’arbre quaternaire peut être levée. L’utilisation des dérivées sur un maillage en facettes passe par l’emploi d’interpolants polynômiaux avec suffisamment de degrés de liberté (cubiques ou de degré supérieur). L’inconvénient de ces fonctions est de ne pas être très stables, ce qui pose des problèmes si le calcul de la dérivée est entâché d’erreurs. L’approche choisie l’a été pour sa plus grande stabilité.

4.1 Échantillonnage de l’illumination

4.1.1 Placement des échantillons

Arbres d’échantillons

On considère dans la suite des surfaces paramétriques $M(u, v) \in \mathbb{R}^3$ définies sur un domaine rectangulaire $[a, b] \times [c, d]$ comme par exemple un quadrilatère ou un carreau spline. A chaque surface, on associe un ou plusieurs arbres quaternaires qui divisent l’espace des paramètres. La méthode peut également être adaptée à des facettes ou carreaux triangulaires en remplaçant les arbres quaternaires par des arbres triquaternaires semblables à ceux qui sont décrits dans [BMSW91].

Chaque noeud de l’arbre représente donc un sous-rectangle de $[a, b] \times [c, d]$. Les points P_i où l’illumination va être échantillonnée sont les images dans l’espace objet de points q_i situés dans chacun de ces sous-rectangles (voir figure 4.1). Contrairement aux cas des maillages en facettes, il n’y a donc pas de partage de

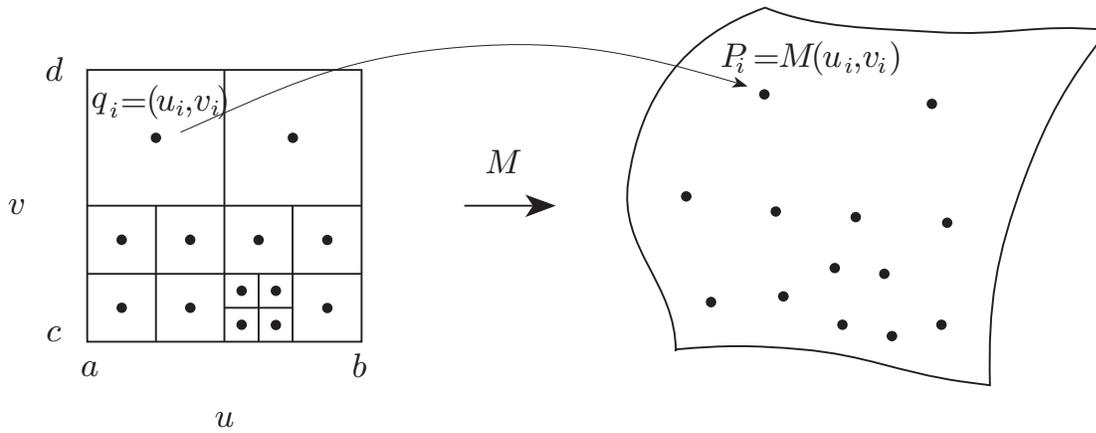


FIG. 4.1 - Placement des échantillons.

sommets entre facettes voisines car chaque P_i se trouve strictement à l'intérieur de l'élément de surface auquel il est associé.

Régularisation de la distribution

Si on place q_i au centre du sous-rectangle, le placement des échantillons est régulier dans le domaine paramétrique mais ne l'est pas forcément dans l'espace 3D. Or le maillage est créé sans information *a priori* sur la fonction de radiosité qui y sera stockée et le placement optimal pour détecter les variations et raffiner est donc régulier dans l'espace 3D. Ce problème peut être résolu en changeant la paramétrisation de la surface pour la rendre plus proche de l'abscisse curviligne. Cette solution serait sans doute nécessaire si l'étape de reconstruction de la radiosité se faisait dans l'espace paramétrique. Nous verrons dans la suite que cette étape est réalisée directement dans l'espace objet. Sachant que le temps de création du maillage est largement dominé par celui des calculs de radiosité qui le suivent, nous proposons ici une solution simple à mettre en œuvre même si son coût n'est pas optimal.

Etant donné un rectangle $[a_i, b_i] \times [c_i, d_i]$ dans le domaine paramétrique, nous voulons y choisir un point $q_i = (u_i, v_i)$ dont l'image P_i sera située sensiblement au centre du carreau image de $[a_i, b_i] \times [c_i, d_i]$ dans \mathbb{R}^3 . Pour cela, on approche la courbe isoparamétrique $v = (c_i + d_i)/2$ par une ligne polygonale formée par les points $M(u_j, (c_i + d_i)/2)$ où les u_j divisent régulièrement l'intervalle $[a_i, b_i]$. On calcule ensuite le point milieu I_u de cette ligne qui se situe, par exemple, sur le segment $(M(u_{j_0}, (c_i + d_i)/2), M(u_{j_0+1}, (c_i + d_i)/2))$. Finalement, u_i est calculée par règle de trois dans l'intervalle $[u_{j_0}, u_{j_0+1}]$. Une opération similaire permet d'obtenir v_i et les longueurs des deux lignes polygonales sont également utilisées pour estimer l'aire de l'élément de surface $M([a_i, b_i] \times [c_i, d_i])$. L'approximation de la longueur des courbes u ou $v = \text{cste}$ peut-être rendue aussi précise que l'on veut en augmentant le nombre de segments de la ligne polygonale si la surface ondule

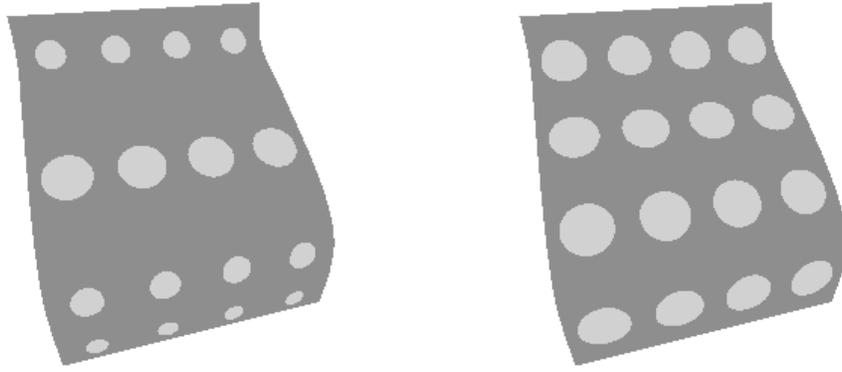


FIG. 4.2 - *Adaptation de la subdivision d'une surface paramétrique pour obtenir un placement régulier des échantillons dans l'espace objet. La surface est mal paramétrée rendant mauvaise la subdivision régulière en (u,v) (à gauche). La méthode décrite permet d'adapter la subdivision dans l'espace objet (à droite). La taille des cercles qui sont centrés sur les échantillons est proportionnelle à l'aire de l'élément de surface associé.*

beaucoup. En pratique, on peut choisir ce nombre quand on place l'échantillon de la racine de l'arbre puis le diviser par deux à chaque fois qu'on descend d'un niveau dans l'arbre pendant le raffinement.

Si le noeud associé au rectangle $[a_i, b_i] \times [c_i, d_i]$ doit être subdivisé par la suite, les quatre sous-rectangles formés seront $[a_i, u_i] \times [c_i, v_i]$, $[a_i, u_i] \times [v_i, d_i]$, $[u_i, b_i] \times [c_i, v_i]$ et $[u_i, b_i] \times [v_i, d_i]$. Ce choix, fait indépendamment des subdivisions des noeuds voisins, est acceptable puisqu'aucun point du maillage n'est partagé entre sous-carreaux voisins. La figure 4.2 montre la distribution plus régulière du maillage obtenue grâce à cette technique.

Le placement adapté des échantillons peut aussi s'appliquer à d'autres objets. Par exemple, nous utilisons le paramétrage de la sphère décrit dans [Dev88] et modifié de manière à répartir régulièrement les échantillons (voir la partie gauche de la figure 4.3). Ce paramétrage divise la sphère en six domaines $[-1, 1] \times [-1, 1]$ en inscrivant un cube à l'intérieur. Chaque point est paramétré par les coordonnées de sa projection centrale sur l'une des six faces du cube. La distribution d'échantillons qui en résulte sur la sphère n'est pas régulière et présente des accumulations vers les coins du cube. Pour changer cette distribution sur les faces du cube, on a calculé l'aire $\mathcal{A}_{a,b}$ sur la sphère de l'image d'un rectangle $[a, b] \times [a, b]$ situé sur une face du cube :

$$\mathcal{A}_{a,b} = \arctan \frac{a^2}{\sqrt{1+2a^2}} + \arctan \frac{b^2}{\sqrt{1+2b^2}} - 2 \arctan \frac{ab}{\sqrt{1+a^2+b^2}}$$

On utilise cette formule pour découper récursivement les rectangles $[a, b] \times [a, b]$

situés sur la diagonale du carré de paramètres en calculant numériquement m tel que $\mathcal{A}_{a,m} = \mathcal{A}_{m,b}$. Seules les cellules situées sur la diagonale ont toutes même aire mais en pratique la subdivision obtenue est régulière sur la sphère (voir la partie droite de la figure 4.3). Cette technique est à rapprocher de celles employées pour obtenir une distribution aléatoire non uniforme à partir d'échantillons uniformément répartis en inversant la fonction de distribution [Shi90].

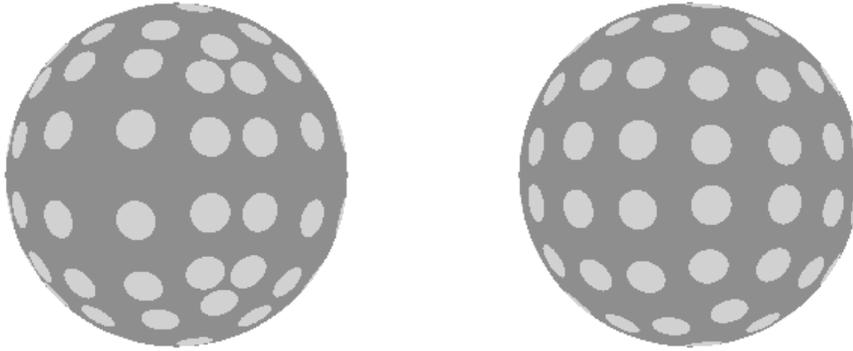


FIG. 4.3 - *Distribution régulière dans l'espace paramétrique (à gauche) et dans l'espace objet (à droite) pour la sphère.*

Échantillonnage initial

La densité d'échantillonnage ou de manière équivalente l'aire des éléments de surface de départ est choisie par l'utilisateur. Le maillage est ensuite initialisé par création d'arbres complets sur les surfaces. Les techniques de régularisation de la distribution présentées assurent non seulement que le maillage est presque uniforme sur chaque surface mais également que sa densité est dans un rapport au plus 4 avec celle fixée. À partir du maillage régulier de départ, le raffinement se fait par subdivision des feuilles de l'arbre selon des critères que nous détaillons maintenant.

4.1.2 Critères pour le raffinement

Le raffinement des maillages de polygones se fait habituellement en subdivisant les facettes pour lesquelles l'écart entre les radiosités des sommets dépasse un seuil fixé par l'utilisateur [CGIB86, BMSW91]. Ce critère fournit une meilleure représentation des bords d'ombre et de pénombre en concentrant les échantillons dans les zones où la radiosité varie beaucoup. Cependant, il a l'inconvénient de subdiviser finement et de manière uniforme l'ensemble des pénombres alors qu'une classification plus fine permettrait d'utiliser moins d'échantillons. Bien qu'il y ait

des discontinuités à l'intérieur des zones de pénombre, la fonction d'illumination y est lisse pour l'essentiel.

Ce sont, comme nous l'avons vu au §2.3, les discontinuités qui, à défaut d'être introduites explicitement, réclament une division plus fine. Le critère de raffinement que nous proposons fait en sorte que les échantillons soient concentrés le long de celles qui sont facilement décelables, c'est-à-dire les transitions entre ombre, pénombre et pleine illumination. Pour cela, le nombre de rayons interceptés dans le calcul du vecteur de forme permet d'étiqueter les échantillons suivant qu'ils voient la source complètement (pleine illumination), en partie (pénombre) ou bien pas du tout (ombre). On parcourt ensuite le maillage pour subdiviser les paires d'échantillons voisins étiquetés différemment. Avec ce choix, les pénombres ne sont pas aussi finement subdivisées qu'avec un simple test d'écart de radiosité entre voisins. Un test supplémentaire est ajouté dans les zones de pleine illumination. Il consiste à comparer la différence d'intensité entre deux échantillons voisins avec la valeur déduite de la dérivée et de la distance entre les échantillons.

Ce critère, fonctionne bien en pratique, mais présente encore quelques défauts :

- les discontinuités situées à l'intérieur des pénombres ne sont pas détectées car l'étiquetage à trois valeurs utilisé est trop grossier ;
- les discontinuités ne sont pas trouvées à leur emplacement réel car les rayons utilisés pour tester la visibilité sont lancés vers des points situés strictement à l'intérieur de la source et non sur son bord.

Ces points peuvent être améliorés à condition de changer la méthode employée pour calculer les vecteurs de forme. Par exemple, si des rayons sont lancés vers les sommets d'une source polygonale, certaines des discontinuités de type EV peuvent être facilement détectées. Ce point est évoqué dans le chapitre 6 consacré à l'intégration de contour approché.

Dans la technique de maillage décrite par Baum *et al.*, une contrainte supplémentaire est imposée aux arbres triquaternaires. On interdit une différence de profondeur strictement supérieure à un entre deux facettes voisines (on parle alors d'arbre restreint). Cette technique permet de retrouver et d'éliminer facilement du maillage les sommets en T pour éviter les discontinuités de la fonction de radiosité reconstruite. L'inconvénient majeur de cette solution est qu'une subdivision en entraîne de nombreuses autres en cascade dans des zones où elles sont en général inutiles. Avec la méthode de reconstruction utilisée dans notre cas et décrite dans la section suivante, ce problème n'existe pas et l'arbre quaternaire n'est pas restreint. Nous verrons cependant que pour obtenir ce résultat, le choix des fonctions de base utilisées pour la reconstruction doit être fait avec soin (§4.2.3).

4.2 Reconstruction de la radiosit 

Apr s convergence de l'algorithme de radiosit  progressive, nous disposons d'un certain nombre de points  chantillons P_i pour lesquels la valeur de radiosit  B_i est connue. Pour calculer des images de la sc ne, il faut pouvoir d terminer la radiosit  en un point quelconque P   partir des B_i . Cette op ration est en fait un probl me d'interpolation des valeurs aux  chantillons et de nombreuses m thodes ont  t  d velopp es pour le r soudre [Fra82]. Nous avons choisi au d part de ne pas d pendre d'une division en facettes ou d'une triangulation du domaine ce qui  limine, par exemple, l'emploi de fonctions de base de type  l ments finis. Nous utilisons une m thode qui associe   chaque valeur B_i une fonction de pond ration, $w_i(\Delta_i P)$, d pendant de la distance $\Delta_i P$ de P_i   P , et qui calcule la valeur en tout point par moyenne pond r e des B_i . Ce type de m thode ne fournit pas un r sultat interpolant les valeurs aux  chantillons   moins de consid rer les coefficients des w_i comme des inconnues et de r soudre ensuite un syst me lin aire pour forcer l' galit    B_i en chaque P_i . Nous verrons au chapitre suivant que le choix d'une fonction $w_i(\Delta_i P)$ d croissant assez vite de sa valeur maximale prise en 0 permet d'obtenir malgr  tout une bonne approximation. Par exemple, Ward *et al.* utilisent comme pond ration une fonction de l'inverse de la distance et de l' cart de normale [WH92]  tendant ainsi une m thode d velopp e par Shepard [She68]. A cause du poids tr s fort de B_i en P_i , ces fonctions ont l'inconv nient de reconstruire une fonction de radiosit  pratiquement constante (plate) dans un voisinage de P_i . Les d riv es aux points P_i peuvent  tre utilis es pour  liminer ce d faut [WH92].

Dans ce travail, nous avons utilis  des fonctions de base gaussiennes   support tronqu , qui permettent de reconstruire une fonction de radiosit  sans discontinuit s. Nous verrons que ce choix permet un traitement satisfaisant de la subdivision adaptative du maillage et fournit de bons r sultats visuels.

Nous pr sentons ici deux m thodes diff rentes (filtres elliptiques et extrapolation) pour calculer la valeur de la radiosit  en un point quelconque d'une surface, la premi re a  t  d crite dans [Ved92] alors que la seconde est une variante qui s'inspire de [WH92]. Dans les deux cas, la reconstruction en un point P se fait en utilisant la formule :

$$\vec{B}(P) = \frac{\sum_i w_i(\Delta_i P) \vec{B}_i}{\sum_i w_i(\Delta_i P)}, \quad (4.1)$$

o  \vec{B} est le vecteur de radiosit , produit de la radiosit  scalaire par le vecteur de forme et w_i est une fonction gaussienne de la distance   P_i . L'utilisation du vecteur de forme permet, au moins pour l' clairage direct, de ne pas avoir   introduire la diff rence de normale entre P et P_i dans la pond ration contrairement   [WH92]. Le choix de la variance de la gaussienne est un compromis   faire entre antialiasage et flou des bords d'ombre [MN88]. Nous y reviendrons dans la section 4.2.3 consacr e au choix des fonctions de pond ration des sous- chantillons

au moment de la subdivision.

4.2.1 Filtres elliptiques

Cette première méthode est applicable même si on ne connaît que la direction \vec{d}_i de variation maximale de la radiosité pour chaque échantillon P_i . Elle consiste à déformer la fonction de poids w_i associée à P_i suivant cette direction de la manière suivante. Soit \vec{d}_i^\perp la direction orthogonale à \vec{d}_i , la fonction déformée w_i^d se calcule comme suit :

$$w_i^d(\Delta_i P) = w_i(\alpha \vec{\Delta}_i P \cdot \vec{d}_i + (1 - \alpha) \vec{\Delta}_i P \cdot \vec{d}_i^\perp),$$

où $1/2 < \alpha < 1$. Ainsi, la gaussienne est comprimée suivant \vec{d}_i et étirée suivant \vec{d}_i^\perp . Le poids total de P_i , c'est-à-dire l'intégrale de w_i , est inchangé. Lors de la reconstruction, l'échantillon aura ainsi un poids moins important pour les points situés dans la direction de plus forte variation de la radiosité.

Cette méthode permet d'obtenir de meilleures fonctions reconstruites qu'avec les filtres circulaires avec peu d'informations sur le gradient. Elle possède cependant deux limitations importantes. D'une part, elle ne permet pas vraiment de tenir compte de l'amplitude du gradient si on l'a calculée. En effet le choix du paramètre α est intuitif et sa valeur n'a pas de signification concrète. De plus, les supports des w_i étant tronqués, l'intervalle des valeurs possibles de α est limité à $[1/2, 1]$ car il ne doit pas y avoir de trous entre les supports de fonctions d'échantillons voisins. Dans les tests de cette méthode qui ont été présentés dans [Ved92], le paramètre α était fixée à une même valeur partout.

L'autre défaut des filtres elliptiques est l'impossibilité d'additionner simplement les illuminations issues de plusieurs sources différentes dans un même échantillon. Chaque source doit en effet posséder une fonction de poids déformée w_i^d différente à chaque échantillon. L'accumulation de l'éclairage ne peut se faire qu'au moment de la reconstruction. En chaque point P , les moyennes 4.1 sont calculées séparément pour les différentes sources puis additionnées à la radiosité totale. Le coût de stockage et de reconstruction est donc augmenté avec chaque nouvelle source et la technique ne peut être appliquée en pratique qu'aux sources primaires. Une autre technique de reconstruction inspirée de [WH92] permet de résoudre ces problèmes.

4.2.2 Filtres circulaires et extrapolation

La reconstruction est une moyenne pondérée $\sum_i w_i B_i / \sum_i w_i$ des valeurs des échantillons. Dans la méthode précédente, w_i était modifiée pour tenir compte des informations supplémentaires sur le gradient. On a vu les problèmes posés en particulier par la multiplication des fonctions de poids w_i^d indépendantes pour chaque P_i qui oblige à séparer les contributions des différentes sources jusqu'à

une  tape avanc e du calcul. La solution consiste   modifier la valeur B_i de l' chantillon plut t que son poids w_i . Pour cela, on utilise un d veloppement limit  au premier ordre pour calculer la valeur de l' chantillon P_i extrapol e en P . Dans la moyenne pond r e, on remplace donc \vec{B}_i par :

$$\vec{B}_i + \left(\partial_j \vec{B}_i\right) \vec{\Delta}_i P, \quad (4.2)$$

o  $\left(\partial_j \vec{B}_i\right)$ est la matrice des d riv es partielles de \vec{B}_i (Jacobien) calcul e en utilisant l'algorithme d crit au chapitre pr c dent. L' quation 4.1 devient donc :

$$\vec{B}(P) = \frac{\sum_i w_i(\Delta_i P) \left(\vec{B}_i + \left(\partial_j \vec{B}_i\right) \vec{\Delta}_i P\right)}{\sum_i w_i(\Delta_i P)}, \quad (4.3)$$

Cette formulation a l'avantage de permettre d'accumuler les contributions de plusieurs sources pour un m me  chantillon puisque la d rivation partielle est un op rateur lin aire et qu'on a donc :

$$\sum_i \left(\vec{B}_i + \left(\partial_j \vec{B}_i\right) \vec{\Delta}_i P\right) = \sum_i \vec{B}_i + \left(\partial_j \sum_i \vec{B}_i\right) \vec{\Delta}_i P. \quad (4.4)$$

4.2.3 Fonctions de pond ration et subdivision

Le probl me de l'addition de plusieurs sources n'est pas encore compl tement r solu. En effet, des sources diff rentes cr ent des ombres et donc, puisque le maillage est adaptatif, des distributions d' chantillons diff rentes. Concr tement, cela signifie que la contribution d'une lampe peut  tre stock e dans un noeud de l'arbre alors que celle de la suivante est stock e dans ses quatre sous-noeuds. L'algorithme de reconstruction ne peut pas utiliser directement de telles donn es car chaque  chantillon doit repr senter l' clairage total pour  tre pris en compte dans la formule 4.1. La seule solution consiste donc   stocker toutes les lampes   la r solution la plus fine localement, c'est- -dire aux feuilles de l'arbre. Cela ne veut pas dire que l'illumination due   chaque lampe sera calcul e   cette r solution ce qui serait contraire aux objectifs du maillage adaptatif et n cessiterait des retours en arri re dans le calcul de radiosit  progressive. On est donc amen    r soudre le probl me suivant : comment distribuer la contribution d'un noeud dans ses quatre sous-noeuds pour lui conserver un effet semblable dans la formule 4.1.

Les m thodes de maillage utilisant des valeurs aux sommets des facettes rencontrent le m me probl me pour subdiviser une facette contenant l' clairage de lampes trait es auparavant. Deux solutions ont  t  propos es dans ce cas : calcul des nouvelles valeurs par interpolation ou recalcul de l' mission des sources pr c dentes sur le nouveau maillage [BMSW91]. A notre avis, la seconde solution, outre qu'elle force   recommencer des calculs d j  effectu s, n'apporte rien. En effet, si l'algorithme de maillage adaptatif a arr t  la subdivision   ce niveau pour les sources pr c dentes, c'est que la repr sentation de la radiosit  ainsi obtenue

était satisfaisante compte tenu du seuil d'erreur fixé au départ. Il suffit donc d'assurer que la radiosité sur les sous-facettes sera identique à la restriction de celle de la facette ce qui est toujours possible pour les schémas d'interpolation polynomiaux.

Dans notre cas, le choix n'est pas aussi simple. Supposons que l'échantillon P_i de radiosité B_i et de poids w_i doit être remplacé par les quatre échantillons P_i^k de radiosités B_i^k et de poids w_i^k . Le problème posé plus haut revient à trouver les B_i^k et w_i^k tels que l'on ait :

$$w_i(\Delta_i P) \left(\vec{B}_i + \left(\partial_j \sum_i \vec{B}_i \right) \vec{\Delta}_i P \right) = \sum_k w_i^k(\Delta_i^k P) \left(\vec{B}_i^k + \left(\partial_j \vec{B}_i^k \right) \vec{\Delta}_i^k P \right) \quad (4.5)$$

et

$$w_i(\Delta_i P) = \sum_k w_i^k(\Delta_i^k P) \quad (4.6)$$

pour tout point P .

Pour les B_i^k et les $\partial_j \vec{B}_i^k$, le choix évident de :

$$\begin{aligned} \vec{B}_i^k &\leftarrow \vec{B}_i + \left(\partial_j \vec{B}_i \right) \vec{\Delta}_i P_i^k, \\ \left(\partial_j \vec{B}_i^k \right) &\leftarrow \left(\partial_j \vec{B}_i \right). \end{aligned}$$

permet de réduire les équations 4.5 et 4.6 à 4.6 seule puisque l'on a alors :

$$\forall k, \quad \vec{B}_i + \left(\partial_j \sum_i \vec{B}_i \right) = \vec{B}_i^k + \left(\partial_j \vec{B}_i^k \right) \vec{\Delta}_i^k P.$$

La décomposition exacte d'une gaussienne en quatre sous-gaussiennes étant impossible, nous allons chercher une solution approchée dans le cas d'un échantillon situé dans une cellule plane et carrée. Les points P_i et P_i^k sont alors fixés et nous pouvons en outre supposer que les fonctions w sont identiques à une profondeur donnée dans l'arbre d'échantillons. A la profondeur n dans l'arbre quaternaire, la fonction de poids w est déterminée par sa variance σ_n^2 et par son coefficient de normalisation λ_n .

Nous pouvons donc placer deux exigences différentes sur les gaussiennes des sous-échantillons placés en P_i^k :

- 1° pour la source dont on est en train de calculer l'illumination, le recouvrement des supports des fonctions de poids entre échantillons voisins doit être conservé à tous les niveaux de l'arbre pour obtenir une bonne reconstruction ;
- 2° pour les contributions des sources précédentes, il faut que la somme des quatre nouvelles gaussiennes centrées en P_i^k soit aussi proche que possible de l'ancienne gaussienne centrée en P_i .

Ces deux contraintes ne peuvent pas être réalisées simultanément (voir la figure 4.6 page 58).

Dans les deux cas, on impose que le poids total soit conservé par la subdivision ce qui s'écrit :

$$\frac{\lambda_n}{\sigma_n^2} = 4 \frac{\lambda_{n+1}}{\sigma_{n+1}^2}. \quad (4.7)$$

Pour le 1° , la conservation du recouvrement impose en plus que $\sigma_{n+1}^2 = 4\sigma_n^2$, puisque le raffinement divise la taille des cellules par deux, ce qui donne finalement :

$$\begin{aligned} \sigma_{n+1} &= 2\sigma_n, \\ \lambda_{n+1} &= \lambda_n. \end{aligned} \quad (4.8)$$

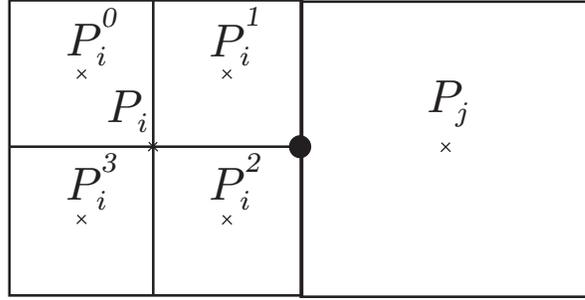


FIG. 4.4 - Point choisi pour imposer la coïncidence entre la fonction w_i de P_i et la somme des fonctions w_i^k des sous-échantillons P_i^k .

Pour le 2° , la décomposition exacte étant impossible, nous imposons comme contrainte supplémentaire que la somme des quatre gaussiennes coïncide avec la gaussienne de départ en un point particulier, le milieu du segment joignant P_i à un de ces quatre voisins du même niveau (voir figure 4.4). Cette contrainte permet de donner à $\sum_k w_i^k$ un poids semblable à celui de w_i vis à vis des échantillons voisins et s'écrit :

$$2l_{n+1} \left(e^{-5\sigma_n^2 l_{n+1}/2} + e^{-\sigma_n^2 l_{n+1}/2} \right) = e^{-\sigma_n^2/4}, \quad (4.9)$$

où on a posé $l_{n+1} = \lambda_{n+1}/\lambda_n$. Cette équation peut être résolue numériquement. La suite des valeurs de σ_n^2 et λ_n dans les deux cas est donnée dans le tableau 4.1.

Le premier choix diminue le flou des fonctions reconstruites mais oblige à restreindre l'arbre quaternaire. En effet, le recouvrement entre fonctions de pondération voisines est bon pour un niveau donné dans l'arbre mais mauvais entre niveau différents ce qui oblige à restreindre ces différences (voir la figure 4.7 page 59).

		0	1	2	3
1°	σ_n^2	2.0	8.0	32.0	128.0
	λ_n	1.0	1.0	1.0	1.0
2°	σ_n^2	2.0	2.70	3.97	9.19
	λ_n	1.0	0.338	0.124	0.072

TAB. 4.1 - Table des valeurs de σ_n^2 et λ_n pour les deux choix de subdivision et différents niveaux dans l'arbre.

Le deuxième choix permet de se passer de la restriction de l'arbre quaternaire sans introduire d'artefacts à la limite entre niveaux différents. Cependant la résolution des bords d'ombre est moins bonne car au niveau le plus fin, la valeur de σ_2 utilisée est plus petite et donc le recouvrement entre les w_i d'échantillons voisins plus grand. En pratique, on utilise une solution de compromis qui consiste à choisir :

$$\begin{aligned}\sigma_{n+1} &= \sqrt{3} \sigma_n, \\ \lambda_{n+1} &= 3 \lambda_n/4\end{aligned}\tag{4.10}$$

et qui permet de ne pas imposer de restriction à l'arbre quaternaire tout en conservant une bonne reconstruction des bords d'ombre.

4.3 Composante directe et indirecte de l'illumination

Les fuites de lumière ou d'ombre sont un des défauts qui apparaissent quand on reconstruit l'illumination à partir d'échantillons sans introduire les discontinuités. Ces deux problèmes sont illustrés par la figure 4.5. Il faut noter que l'erreur commise n'est pas plus importante en valeur que dans une autre partie du maillage mais qu'elle est plus gênante car immédiatement visible. Pour une scène faite de polyèdres, Baum *et al.* proposent de résoudre ce problème en détectant les arêtes situées à l'intérieur d'une face pour les introduire explicitement dans le maillage de la scène. L'utilisation d'échantillons mal placés est alors évitée. Les maillages de discontinuités évitent aussi naturellement ces problèmes qui sont surtout sensibles pour les maillages réguliers.

Dans le cas d'un maillage adaptatif, le problème est plus marqué pour l'éclairage indirect que pour l'éclairage direct. En effet, l'algorithme de subdivision raffine le maillage le long des arêtes posant problème car elles correspondent à des limites entre ombre et pleine illumination. L'éclairage indirect est lui obtenu par accumulation de petites contributions qui sont donc calculées à un niveau

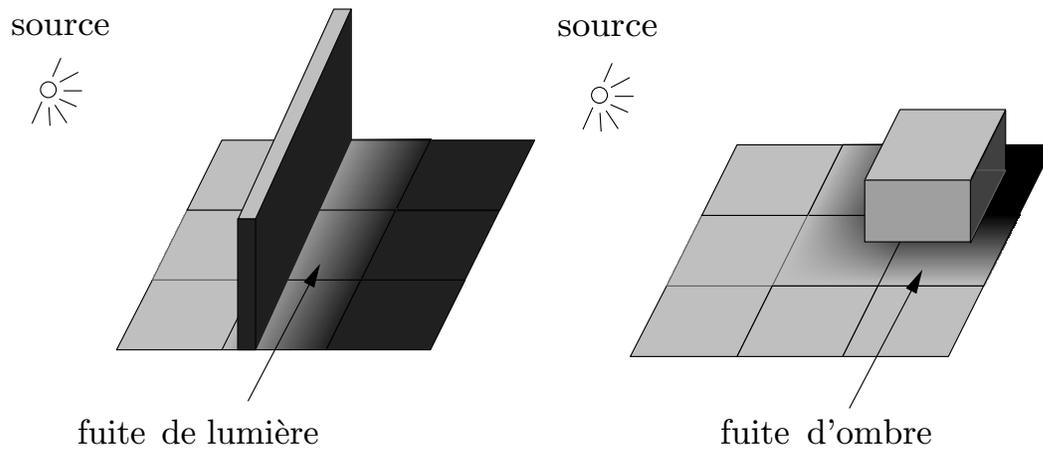


FIG. 4.5 - A gauche, un exemple de fuite de lumière. Des échantillons éclairés sont utilisés pour reconstruire la radiosité de facettes qui traversent le mur. De la lumière apparaît donc dans une zone qui devrait être à l'ombre. A droite, un exemple de fuite d'ombre. Les échantillons cachés par la boîte ont une radiosité nulle qui se propage par interpolation autour de la boîte.

plus élevé de la subdivision puis descendues par interpolation jusqu'au feuilles. Au total, l'effet de fuite d'ombre pour l'éclairage indirect est important et n'est pas réduit par le raffinement du maillage.

Supposons maintenant que l'éclairage indirect soit calculé sur un maillage qui n'est jamais raffiné. Un échantillon situé sous un objet ne recevra jamais d'énergie durant tout le calcul de radiosité progressive et pourra donc être facilement repéré à la fin. Dans la méthode de reconstruction présentée dans la section précédente, il suffit alors d'exclure cet échantillon de la moyenne calculée par la formule 4.1 pour éliminer les fuites d'ombre. Nous proposons donc de stocker la composante indirecte de l'éclairage, indépendamment de sa composante directe, dans un maillage fixé au départ. Les tests du chapitre suivant montrent que l'éclairage indirect varie lentement et peut être représenté avec peu d'échantillons. De plus, les zones réclamant un maillage plus fin peuvent être déterminées *a priori*. Nous présentons aussi dans ce chapitre des résultats montrant l'élimination des fuites d'ombres dans le cas d'objets empilés (§5.3.3).

4.4 Radiosité et surfaces courbes

L'éclairage des surfaces courbes est *a priori* plus complexe que celui des polygones. En effet, il dépend également des variations de normale. Si c'est la radiosité scalaire qui est échantillonnée, la variation de normale impose en général une fréquence d'échantillonnage plus élevée que ce qu'elle serait pour un même éclairage d'une surface plane, avec pour conséquence une augmentation du temps de calcul

et de la mémoire utilisée. La variation continue de la normale est interpolée entre trop peu de positions clés et ce sous-échantillonnage fait apparaître des bandes de Mach particulièrement autour des crêtes et des vallées de la surface.

Si on se place sur une surface pour laquelle la source est toujours au dessus de l'horizon, l'emploi du vecteur de forme permet de résoudre ce problème. Celui-ci, si il varie peu, peut en effet être représenté précisément avec peu d'échantillons (voir le bas de la figure 4.8). L'échantillonnage fin de la normale est reporté à l'étape de reconstruction d'où un gain en temps de calcul et en taille mémoire nécessaire au stockage. Même si l'emploi d'un vecteur en remplacement d'un scalaire triple la taille de l'échantillon lumineux, on obtient un gain net en occupation mémoire. Ce gain est permis par la diminution du nombre total d'échantillons et par le fait que l'information lumineuse ne représente qu'une partie de la structure de données complète (qui contient aussi des informations géométriques, des pointeurs, etc.).

Si la source passe sous l'horizon, c'est-à-dire quand l'éclairage est rasant, le produit scalaire par le vecteur de forme ne prend pas en compte le changement progressif de partie visible. Le résultat est alors incorrect car une partie de la source peut se trouver cachée par l'horizon et il est impossible de déterminer à partir du vecteur total quelle proportion se trouve sous l'horizon. Pour une source primaire, la subdivision adaptative réduit l'erreur qui en résulte en raffinant le maillage. Il s'agit d'une zone de pénombre et le traitement est identique à celui d'une surface plane sur laquelle un autre objet projette une ombre.

Pour les sources secondaires, en revanche, qui forment au total l'éclairage indirect, la situation est la même que dans la section précédente. Les contributions individuelles sont trop faibles pour déclencher un raffinement du maillage mais leur accumulation donne un effet non négligeable.

4.5 Conclusion

Nous avons décrit une méthode de maillage pour la radiosité. Elle utilise une distribution d'échantillons sur les surfaces par des arbres quaternaires. La reconstruction de la radiosité sur les surfaces n'est pas liée à un découpage en facettes et se rapproche de la technique utilisée par Greg Ward dans Radiance. Par contre, c'est une méthode complète de maillage pour la radiosité qui permet de représenter de manière adaptative aussi bien la composante directe de l'éclairage que sa composante indirecte. Les points particuliers de cette méthode sont :

- l'utilisation des dérivées de la radiosité pour améliorer la fonction reconstruite ;
- un choix des fonctions de pondération qui permet de lever la contrainte de restriction de l'arbre d'échantillons ;

- l'utilisation du vecteur de radiosit  sur les surfaces courbes pour  viter un sur chantillonnage inutile.

Dans le chapitre 5, nous pr sentons une  valuation compl te de la m thode de maillage telle qu'elle a  t  d crite. La pr cision de calcul des composantes directe et indirecte de l'illumination est examin e ainsi que l'influence de la subdivision adaptative, de l'utilisation des vecteurs de radiosit  et des d riv es sur cette pr cision.

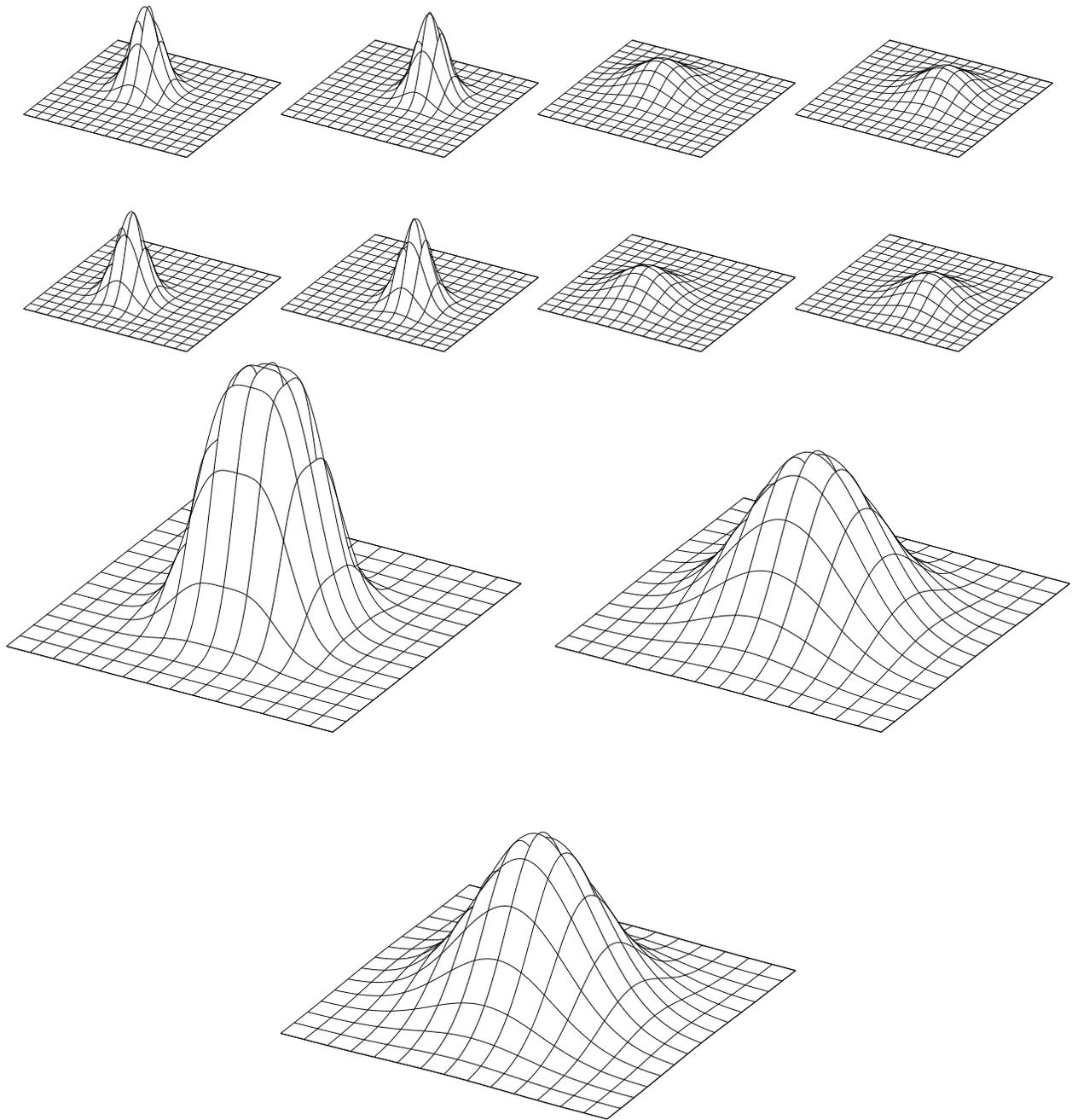


FIG. 4.6 - *Fonctions gaussiennes pour un échantillon et ses quatre sous-échantillons. A gauche, σ^2 est conservé à l'échelle, c'est-à-dire que $\sigma_{n+1}^2 = 4 \sigma_n^2$. A droite, on prend λ_{n+1} solution de l'équation 4.9 et σ_{n+1}^2 tel que le poids total soit conservé.*

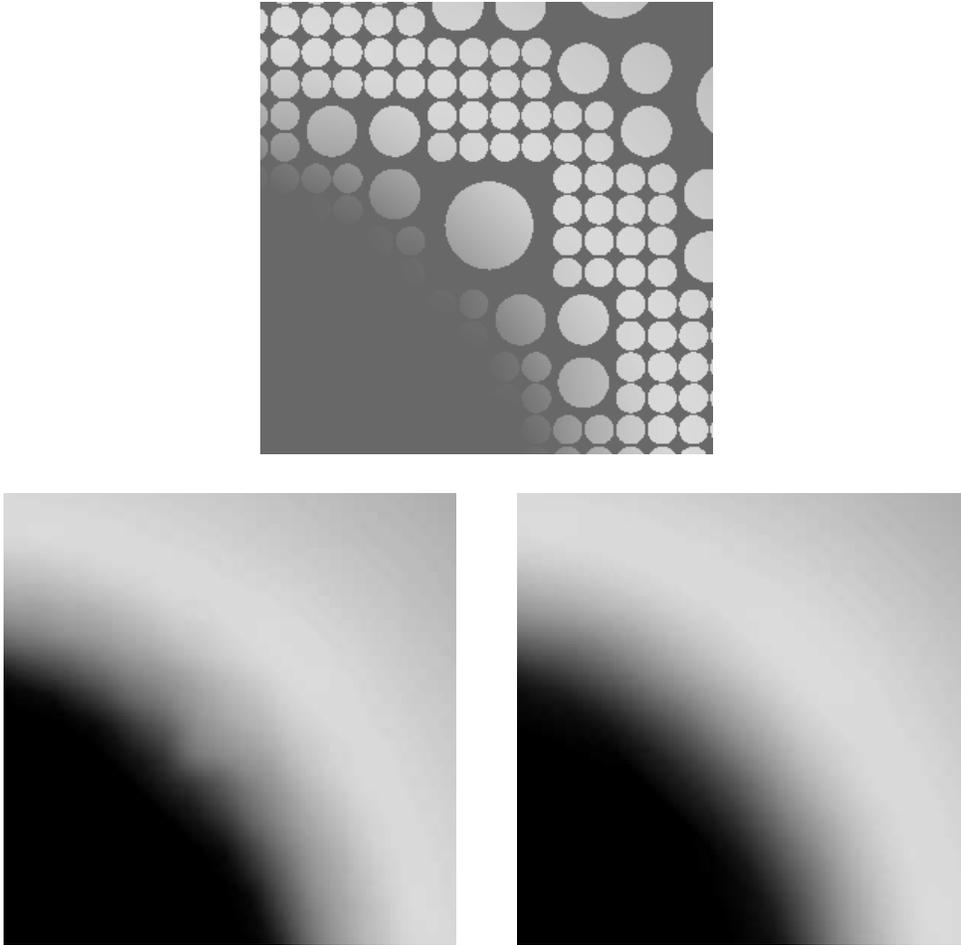


FIG. 4.7 - Défaut de reconstruction dû à un trop grand écart de niveau entre échantillons voisins. En haut, la distribution d'échantillons, on note que la propriété de restriction de l'arbre quaternaire est violée par l'échantillon central. À gauche, la reconstruction obtenue en faisant le choix $\sigma_{n+1}^2 = 4 \sigma_n^2$ et à droite la reconstruction obtenue en faisant le compromis $\sigma_{n+1}^2 = 3 \sigma_n^2$.

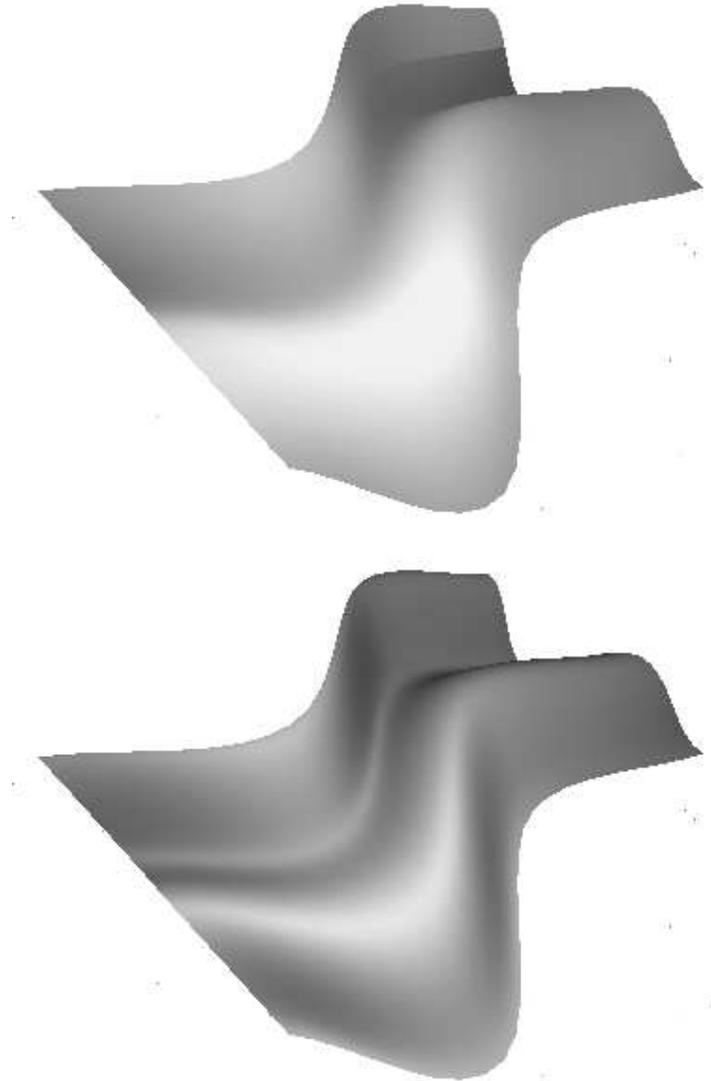


FIG. 4.8 - *Utilisation du vecteur de forme pour reconstruire l'illumination sur les surfaces courbes. Les deux images ont été calculées avec le même nombre d'échantillons (16). En haut, on utilise la radiosité scalaire et la densité d'échantillonnage est trop faible. En bas, on utilise la radiosité vectorielle et on obtient une représentation de la radiosité qui ne peut pas être distinguée de l'image où la radiosité est recalculée à chaque pixel. Cette image de référence n'est pas montrée pour cette raison.*

Chapitre 5

Tests de la radiosité reconstruite

Dans ce chapitre, nous présentons une série de tests de la qualité de simulation obtenue en utilisant la méthode de maillage pour la radiosité proposée au chapitre précédent. L'éclairage direct des sources primaires et l'éclairage indirect par les multiples réflexions de la lumière dans la scène sont traités dans deux parties différentes. Pour ce qui concerne l'éclairage direct, nous nous intéressons particulièrement à l'effet de la subdivision adaptative et de l'utilisation des dérivées sur la précision de la fonction reconstruite. Pour l'éclairage indirect, nous examinons le comportement de la composante indirecte, le problème des fuites d'ombre et l'influence de la finesse du maillage de l'éclairage direct sur la qualité de l'éclairage indirect. Nous commençons par présenter, de manière générale, les méthodes utilisées pour évaluer la validité de l'éclairage calculé.

5.1 Méthodes de test

Dans la suite, nous utilisons deux scènes pour tester la précision de l'éclairage stocké dans le maillage. La première est une «boîte de Cornell», c'est-à-dire l'intérieur d'un cube gris avec deux faces opposées de couleurs bleue et rouge. Contrairement à la scène utilisée à l'origine dans [GTGB84], la lumière ne provient pas de l'une des faces du cube mais d'une source carrée située au plafond. Les effets de coloration par réflexion en rouge ou bleu sont particulièrement faciles à observer sur les faces grises du cube. La deuxième scène est schématisée à la figure 5.1 et éclairée par la source carrée tracée en gris. Cette scène, bien que simple, permet d'étudier à la fois :

- le comportement de la fonction B dans les coins ;
- les problèmes de fuite d'ombre grâce aux objets posés au sol ou appuyés au mur ;
- l'illumination d'une surface courbe (sphère).

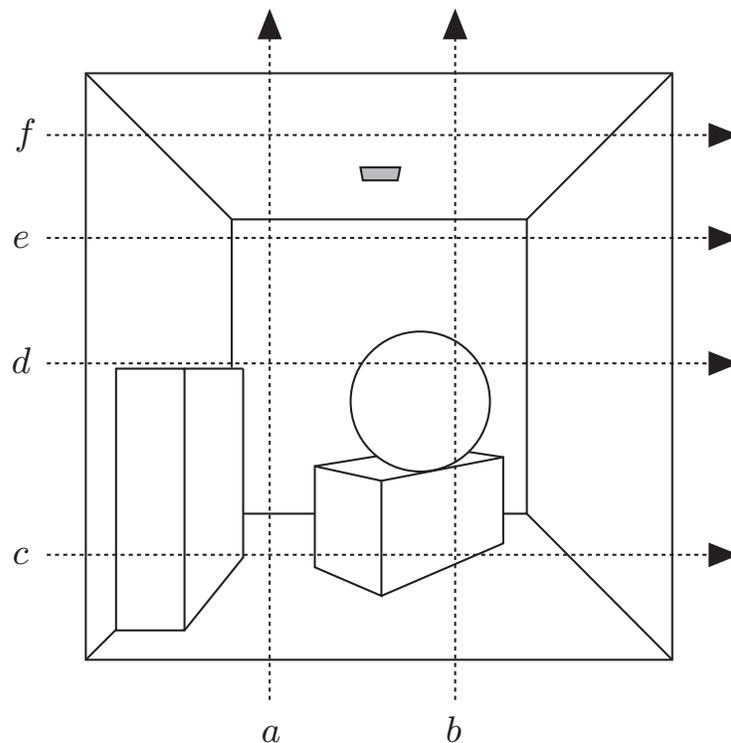


FIG. 5.1 - Scène utilisée pour les tests de cohérence.

La comparaison des niveaux d'erreur entre les deux scènes pour une même densité du maillage permet d'évaluer l'influence des obstructions sur la précision des calculs.

Pour évaluer la justesse de la fonction de radiosité B obtenue par reconstruction à partir d'échantillons d'illumination (méthode décrite au chapitre précédent), on la compare à une autre fonction B' . Le calcul de cette seconde fonction dépend de la nature, directe ou indirecte, de l'éclairage considéré. Dans le premier cas, B' est une valeur très précise de l'éclairage direct calculée pour chaque point où la comparaison est effectuée avec la méthode exposée au chapitre 6. Dans le second cas, l'approche est un peu différente car B' n'est plus une fonction de référence, c'est-à-dire une fonction calculée très précisément et ainsi supposée proche de la solution réelle. B' est obtenue par intégration de la lumière incidente dans l'hémisphère située au dessus du plan tangent à la surface au point considéré. Comme la valeur de l'intensité incidente est elle aussi reconstruite à partir des échantillons d'illumination, c'est en fait la *cohérence* de la solution qui est mesurée. Les détails d'implémentation de chaque méthode sont donnés dans les parties correspondantes.

Le calcul de la fonction B' , surtout dans le cas de l'éclairage indirect où il faut intégrer l'énergie incidente dans l'hémisphère, est coûteux. Il faut donc choisir un ensemble réduit de points sur les surfaces de la scène où la différence entre B et

B' va être mesurée. Nous avons choisi de calculer une vue de la scène par lancer de rayons à basse résolution (128×128) et d'effectuer ces tests pour chaque pixel de l'image ainsi produite. Cette technique fournit à la fois une valeur moyenne et une «image» de l'erreur qui permet de situer les zones critiques dans la scène. Les écarts moyens obtenus, même si ils donnent une indication sur la qualité d'une solution donnée, ne sont pas de vraies moyennes sur l'ensemble des surfaces de la scène. Nous ne les utiliserons que pour comparer plusieurs solutions obtenues avec des paramètres différents (densité du maillage, subdivision adaptative ou non, etc.). Nous effectuons aussi des tests le long de lignes et de colonnes de pixels dans des images de plus grande résolution. On obtient ainsi des profils d'illumination qui permettent d'observer le comportement précis de la radiosité dans des zones particulières de la scène, en s'approchant des coins ou bien sur la sphère par exemple.

Dans la seconde scène, quelques lignes et colonnes particulières ont été choisies pour le calcul de profils d'éclairage. Elles sont étiquetées de a à f et figurent sur le schéma de la scène 5.1. Les flèches indiquent le sens utilisé pour les courbes d'illumination données aux figures 5.6, 5.7 et 5.8.

Dans la suite, on utilise, pour chaque scène, quatre maillages réguliers désignés α , β , γ et δ avec des densités d'échantillonnage croissantes. Ces maillages sont obtenus comme arbres complets et à cause de l'arité des arbres utilisés, le nombre total d'échantillons est approximativement multiplié par 4 quand on passe d'un maillage au suivant. Pour les sections utilisant la subdivision adaptative, le maillage régulier de base est β ou γ . Plusieurs maillages adaptatifs sont obtenus en autorisant des profondeurs maximales de subdivision croissantes. Les maillages correspondants sont notés $\beta + i$, où i désigne le nombre de niveaux supplémentaires de subdivision autorisés. Le tableau 5.1 donne le nombre total d'échantillons pour chacun des maillages utilisés (maillages réguliers pour les deux scènes et adaptatifs pour la seconde uniquement). La figure 5.5, page 74, montre deux exemples de maillages de la deuxième scène. Ils ont été obtenus en tronquant suffisamment les fonctions de poids des échantillons pour qu'elles ne se recouvrent pas.

5.2 Eclairage direct

5.2.1 Calcul de la référence

L'éclairage direct peut être calculé précisément en chaque point pour comparaison avec la valeur reconstruite d'après le maillage. Nous utilisons pour cela la méthode d'intégration de contour qui donne des résultats exacts hors des pénombres. En cas d'obstruction, la technique décrite au chapitre 6 est employée et pour les tests de ce chapitre, des valeurs élevées des paramètres de la méthode (nombre de rayons primaires et de pas de dichotomie) sont choisies.

Maillage	α	β	γ	δ
Scène 1	100	388	1620	6228
Scène 2	135	477	1961	7577

Maillage	β	$\beta + 1$	$\beta + 2$	$\beta + 3$	γ	$\gamma + 1$	$\gamma + 2$	$\gamma + 3$
Scène 2	135	897	1692	3552	1961	2768	4598	8846

TAB. 5.1 - Taille des maillages utilisés pour les deux scènes de test en nombre total d'échantillons.

5.2.2 Subdivision adaptative

Dans cette section, nous nous intéressons aux effets de la subdivision adaptative dans les ombres et pénombres sur la précision de l'éclairage direct. Deux images de la seconde scène ont été calculées pour montrer la meilleure représentation des ombres obtenue grâce au raffinement adaptatif (figure 5.4, page 73).

La figure 5.6, page 74, donne les profils d'illumination à la traversée d'une pénombre pour des maillages adaptatifs de plus en plus fins. On remarque que les erreurs se situent bien aux abords des discontinuités et non dans la pénombre elle-même.

5.2.3 Dérivées de l'illumination

Cette section présente des résultats sur l'utilisation des dérivées de la radiosité lors de la reconstruction. La méthode utilisée est celle qui est fondée sur des filtres de reconstruction circulaires et un développement limité à l'ordre un de la radiosité au voisinage des échantillons (§4.2.2). Dans le tableau 5.2, nous donnons les écarts entre la solution de référence et les solutions reconstruites avec et sans utilisation des dérivées pour une série de maillages adaptatifs de plus en plus fins. L'écart moyen à la solution de référence est toujours inférieur quand on utilise les dérivées. Les profils d'éclairage de la figure 5.7, page 75, montrent que l'amélioration de la fonction reconstruite se trouve à la fois dans les pénombres et les zones pleinement illuminées alors que la section précédente a montré que la subdivision adaptative concerne surtout les pénombres. Dans les zones de pleine illumination, la fonction reconstruite en utilisant les dérivées est indiscernable de la fonction de référence à l'échelle de la figure. Le gain de précision dû aux dérivées est plus important que celui réalisé en augmentant la profondeur de subdivision adaptative, par exemple on passe de 0.99 à 0.83 en allant de $\beta + 1$ à $\beta + 2$ et de 0.99 à 0.63 en utilisant les dérivées. Cette différence s'accroît avec la finesse du maillage.

Maillage	β	$\beta + 1$	$\beta + 2$	$\beta + 3$
Avec dérivée ($\times 10^{-3}$)	1.06	0.63	0.42	0.34
Sans dérivée ($\times 10^{-3}$)	1.46	0.99	0.83	0.75
Maillage	γ	$\gamma + 1$	$\gamma + 2$	$\gamma + 3$
Avec dérivée ($\times 10^{-3}$)	0.41	0.23	0.16	0.13
Sans dérivée ($\times 10^{-3}$)	0.55	0.41	0.36	0.34

TAB. 5.2 - *Tableau de l'écart moyen à la solution de référence pour des radiosités reconstruites avec et sans utilisation de la dérivée.*

La figure 5.8 page 76 présente l'influence de la dérivée pour la reconstruction d'une pénombre. La version calculée avec la dérivée est clairement plus précise mais présente une bosse à la limite entre pleine illumination et pénombre due à la présence d'un échantillon à fort gradient dans la pénombre et proche de cette limite. L'effet symétrique à la limite entre pénombre et ombre est caché par la mise à zéro des valeurs reconstruites négatives mais apparaîtrait si l'ajout de la contribution d'une autre source décalait toute la courbe vers le haut. La taille de la bosse diminue avec le raffinement du maillage mais déjà, pour le cas présenté, l'excès d'illumination locale n'est que de 3 %. De plus, la reconstruction lisse fait qu'il n'est pas accompagné d'une bande de Mach et n'est pas détectable dans l'image. La méthode de reconstruction proposée est donc efficace d'autant plus que le coût de calcul des dérivées est très réduit. La taille de l'échantillon de radiosité est multiplié par quatre (radiosité et trois dérivées partielles contre radiosité seule) mais on note que le raffinement qui multiplie aussi l'échantillon par quatre amène une diminution moindre de l'erreur.

5.3 Eclairage indirect

5.3.1 Notion de cohérence

Principe

Pour juger de la qualité de l'éclairage indirect, plusieurs méthodes sont envisageables. Le calcul d'une solution exacte étant impossible, on utilise souvent comme référence la solution obtenue avec le maillage le plus fin et la meilleure précision de calcul réalisables en pratique. Les solutions moins précises sont ensuite comparées à cette référence, ce qui fournit une erreur dont on peut observer la variation en fonction des paramètres étudiés. Cette méthode ne permet pas de détecter un éventuel biais de l'algorithme utilisé pour calculer la radiosité puisque la solution testée et la référence en proviennent. Ici, nous prenons une approche différente dans laquelle l'erreur associée à une solution approchée n'est pas ob-

tenue par comparaison avec une solution de référence. Pour définir cette erreur, Nous utilisons l'équation d'équilibre de la lumière dans la scène. En tout point x , on a :

$$B(x) = B_e(x) + \int_S B(x')G(x, x')\text{vis}(x, x')dx' \quad (5.1)$$

où $B(x)$ est la radiosité totale au point x , $B_e(x)$ est l'émission propre, $G(x, x')$ est le facteur de forme différentiel entre x et x' et S est l'union des surfaces de l'environnement. Notons $B_d(x)$ et $B_i(x)$, les radiosités directe et indirecte respectivement. $B_d(x)$ représente la partie de $B(x)$ qui a été transportée directement depuis les sources primaires et $B_i(x)$ celle qui a été réfléchi de manière diffuse au moins une fois avant d'atteindre x . En utilisant ces quantités, on peut réécrire l'équation 5.1 sous la forme :

$$B_d(x) + B_i(x) = \int_S [B_e(x') + B_d(x') + B_i(x')]G(x, x')\text{vis}(x, x')dx', \quad (5.2)$$

ou bien encore puisque le passage par x' représente une réflexion diffuse :

$$B_i(x) = \int_S [B_d(x') + B_i(x')]G(x, x')\text{vis}(x, x')dx', \quad (5.3)$$

La fonction $B = B_e + B_d + B_i$ pour laquelle cette égalité est réalisée en tout point et pour laquelle B_d représente bien l'éclairage direct dû à B_e est une solution exacte du problème d'équilibre. Pour une autre fonction de radiosité définie sur les surfaces de la scène, les deux membres de la formule 5.3 ne sont plus égaux. Dans les tests qui suivent, nous utilisons une mesure de l'écart entre ces deux valeurs pour juger de l'erreur de la solution. Cette méthode revient donc à vérifier la cohérence de la fonction B .

Intégration de la lumière incidente

Pour évaluer la cohérence de l'éclairage indirect de la scène en un point x , on calcule sa valeur de deux façons différentes qui correspondent aux deux membres de l'égalité 5.3.

La première méthode consiste simplement à reconstruire la valeur de B_i à partir des échantillons d'illumination comme décrit au §4.2. L'algorithme de radiosité progressive a pour cela stocké séparément B_d et B_i .

Dans la deuxième méthode, on calcule la lumière incidente au point x . Pour échantillonner l'hémisphère située au dessus du point avec la bonne densité, on utilise un algorithme semblable à celui de Malley dans [Mal88] qui est fondé sur l'équivalent de Nusselt. Les échantillons sont distribués uniformément dans un disque centré en x puis relevés sur l'hémisphère pour fournir une direction dans l'espace. Contrairement à [Mal88], la distribution des échantillons est faite sur une grille régulière puisque l'antialiasage n'est pas important pour l'utilisation qui est faite des résultats. L'emploi de cette distribution rend compte du terme

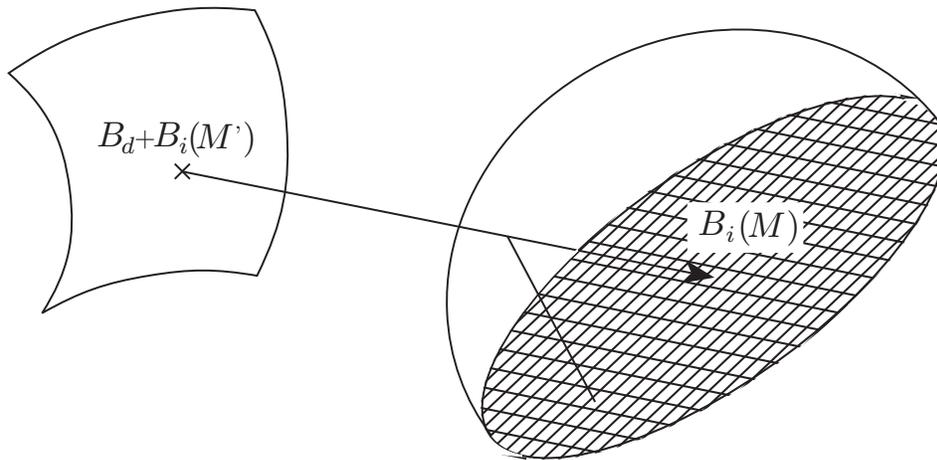


FIG. 5.2 - Calcul de l'éclairage indirect par intégration de la lumière incidente sur l'hémisphère.

$G(x, x')$ de l'intégrale. On lance ensuite des rayons pour déterminer la luminance incidente depuis chacune des directions ainsi obtenue. Si aucun objet n'est intersecté, cette énergie est nulle, sinon elle est égale à l'illumination au premier point d'intersection x' rencontré. Contrairement au cas précédent, on reconstruit en x' la somme de l'éclairage direct B_d et indirect B_i en excluant simplement l'émission propre B_e . En effet, du point de vue de x , la réflexion diffuse en x' assure qu'on a affaire à de l'éclairage indirect. La luminance incidente est ensuite multipliée par la réflectivité en x .

Pour mesurer la cohérence de la radiosité indirecte en un point, on utilise la valeur absolue de l'écart entre les deux valeurs obtenues. La moyenne et le maximum de cette valeur sont calculés sur l'ensemble des pixels de l'image basse résolution comme indiqué précédemment.

Pour choisir la densité d'échantillonnage de l'hémisphère, plusieurs valeurs ont été testées dans quelques cas particuliers. La variation des résultats obtenus pour des densités d'échantillonnage supérieur à 5000 échantillons/hémisphère est négligeable devant les écarts de cohérence mesurés. Les résultats présentés dans la suite ont donc tous été calculés avec 5000 échantillons sur l'hémisphère en chaque point.

Les trois points abordés dans la suite sont :

- le comportement de la valeur reconstruite et de la valeur intégrée de l'éclairage indirecte.
- les «fuites d'ombre» provenant d'empilement d'objets qui ne sont pas pris

en compte explicitement ;

- l'influence de la finesse du calcul de l'éclairage direct sur la cohérence de l'éclairage indirect.

5.3.2 Fonction reconstruite et fonction intégrée

Les figures 5.9 et 5.10, pages 77 et 78, montrent des profils d'éclairage indirect pour les deux scènes tests avec les différents maillages réguliers. A gauche, les courbes correspondent à la valeur reconstruite à partir du maillage, à droite, elles correspondent à la valeur obtenue par intégration de l'illumination incidente dans l'hémisphère. La mesure de qualité de l'éclairage indirect que nous utilisons est l'écart entre ces deux familles de courbes. On remarque que les deux séries de courbes convergent vers la même limite quand la densité du maillage augmente et que les courbes intégrées sont beaucoup plus proches de la limite que les courbes reconstruites pour un même maillage.

Les tableaux 5.3 et 5.4 présentent les écarts moyens et maximaux sur l'image 128×128 pour les quatre maillages réguliers α , β , γ et δ . Le rapport entre écart moyen et valeur moyenne de la radiosité (écart relatif dans le tableau) donne une idée de l'importance de l'erreur. L'image de la figure 5.3 a été calculée avec simplement l'éclairage indirect à partir du maillage δ et montre que B_i varie lentement sur les surfaces de la scène. Ces variations ne sont pas limitées par le maillage qui est beaucoup plus fin. La version de l'éclairage indirect obtenue par intégration confirme ce comportement.

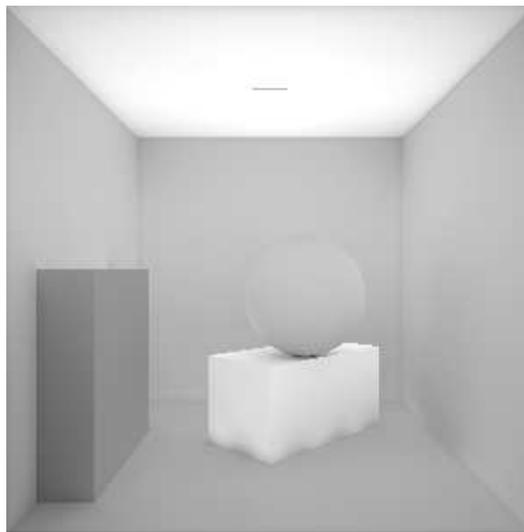


FIG. 5.3 - *Vue de la seconde scène montrant la composante indirecte de l'éclairage seule.*

maillage	nombre d'échantillons	écart moyen 10^{-4}	écart maximal 10^{-2}	écart relatif %
α	100	6.7	3.4	1.73
β	388	3.2	0.3	0.83
γ	1540	2.2	1.2	0.57
δ	6148	2.1	1.2	0.54

TAB. 5.3 - Scène 1: écart entre les deux versions de l'éclairage indirect pour des maillages réguliers de densités croissantes.

maillage	nombre d'échantillons	écart moyen 10^{-4}	écart maximal 10^{-2}	écart relatif %
α	135	27.4	3.6	10.2
β	477	6.6	1.3	2.40
γ	1961	3.5	1.2	1.24
δ	7577	2.8	1.1	1.01

TAB. 5.4 - Scène 2: écart entre les deux versions de l'éclairage indirect pour des maillages réguliers de densités croissantes.

5.3.3 Fuites d'ombre et maillage propre à l'éclairage indirect

Nous donnons ici un exemple d'élimination des fuites d'ombres par la méthode présentée au §4.3. Dans la deuxième scène, les deux parallélépipèdes posés au sol et contre la paroi de gauche cachent plusieurs échantillons d'éclairage indirect. La figure 5.11, page 79, montre deux images de la scène calculées à partir du même maillage de radiosité pour lequel éclairage direct et éclairage indirect ont été stockés séparément. Pour l'image de gauche, tous les échantillons ont été utilisés pour reconstruire la composante indirecte de l'éclairage alors que pour celle de droite, les échantillons contenant une radiosité nulle ont été ignorés. La figure 5.12, page 79, montre les profils de radiosité indirecte le long de la ligne c pour les deux images. Dans la zone où la fuite d'ombre a été éliminée, l'amplitude de l'erreur est comparable à celle du reste de la scène. Cette méthode a bien sûr des limites mais permet de réduire simplement l'effet des fuites d'ombre et améliore la qualité subjective de l'image. Un des problèmes qui ne peut pas être traité simplement sans introduire explicitement les discontinuités de radiosité dans le maillage apparaît dans les courbes de la figure 5.12. L'écart important situé entre les abscisses 90 et 95 est une manifestation de la fuite d'ombre dans le calcul des facteurs de forme. Cette zone, qui correspond à la face avant droite du parallélépipède reçoit une partie de son éclairement du sol situé juste devant. Les éléments de surface dont l'échantillon est caché mais dont une partie dépasse du parallélépipède n'émettent pas de lumière ce qui correspond au déficit observé. Le calcul des discontinuités est donc la solution la plus correcte mais n'est pas forcément applicable pour des objets courbes (colonne cylindrique posée sur le sol).

5.3.4 Influence du maillage de l'éclairage direct

Dans un premier temps, nous présentons les résultats obtenus avec deux échantillonnages réguliers de la scène de densités différentes. Ensuite, ces maillages sont conservés pour l'éclairage indirect mais l'éclairage direct de la source est calculé sur un maillage adaptatif de plus en plus fin. Dans le calcul de facteur de forme par lancer de rayon, on choisit un seuil de subdivision suffisamment petit pour qu'un nombre suffisant de rayons soit lancés pour exploiter le maillage plus fin de l'éclairage direct au moment de sa réémission. On observe dans les deux exemples que si l'erreur diminue quand on raffine l'éclairage direct d'un niveau, elle remonte ensuite régulièrement. Pour expliquer ce comportement paradoxal, nous avons pris comme référence la version intégrée de la composante indirecte pour le maillage le plus fin employé ($\gamma + 3$) puis nous avons calculé les écarts avec toutes les autres versions reconstruites ou intégrées pour les autres maillages. Le

maillage	écart moyen 10^{-4}	écart relatif %	maillage	écart moyen 10^{-4}	écart relatif %
β	6.7	2.4	γ	3.5	1.2
$\beta + 1$	5.4	1.9	$\gamma + 1$	3.6	1.2
$\beta + 2$	5.9	2.1	$\gamma + 2$	3.7	1.3
$\beta + 3$	6.3	2.2	$\gamma + 3$	3.7	1.3

TAB. 5.5 - *Scène 1: écart entre les deux versions de l'éclairage indirect pour des maillages réguliers de densités croissantes.*

résultat de ces comparaisons montre que :

- les versions intégrées de la composante indirecte convergent vers une limite quand on augmente la finesse de représentation de l'éclairage direct;
- les versions reconstruites de la composante indirecte n'évoluent plus après le premier niveau de raffinement de l'éclairage direct.

Ce résultat montre que le maillage de l'éclairage indirecte ne tire pas partie d'un calcul plus précis de l'éclairage direct au delà d'un niveau de raffinement. Nous avons également calculé des histogrammes de l'écart entre fonction intégrée et fonction reconstruite pour les pixels de l'image 128×128 (figure 5.13). Alors que le passage du maillage β au maillage γ plus fin change la forme de l'histogramme et diminue l'écart moyen, le passage à des maillages plus fins de l'éclairage direct $\beta + i$ ou $\gamma + i$ à des effets beaucoup moins marqués.

5.4 Conclusion

Ce chapitre a présenté plusieurs tests des fonctions de radiosit  reconstruites en utilisant la m thode de maillage d crite au chapitre pr c dent. Les r sultats suivants ont  t  obtenus :

- l'utilisation des d riv es pour la reconstruction de la composante directe am liore de mani re importante la fonction reconstruite ;
- certaines fuites d'ombres peuvent  tre  limin es   peu de frais en ignorant les  chantillons de radiosit  nulle lors de la reconstruction ;
- le maillage fin de la composante directe n'est pas n cessaire   un calcul pr cis de la composante indirecte qui est vite limit e par la finesse de son maillage.

Ce dernier point est important pour les algorithmes de radiosité qui recalculent l'éclairage des sources au moment de la production des images car il montre que l'on peut en pratique se contenter d'un maillage non adaptatif.

La méthode d'évaluation de la composante indirecte de l'éclairage proposée pourrait être utilisée pour tester d'autres méthodes de maillage ou de résolution hiérarchique et par exemple étudier l'effet des paramètres de contrôle de ces méthodes sur la précision réelle du résultat.

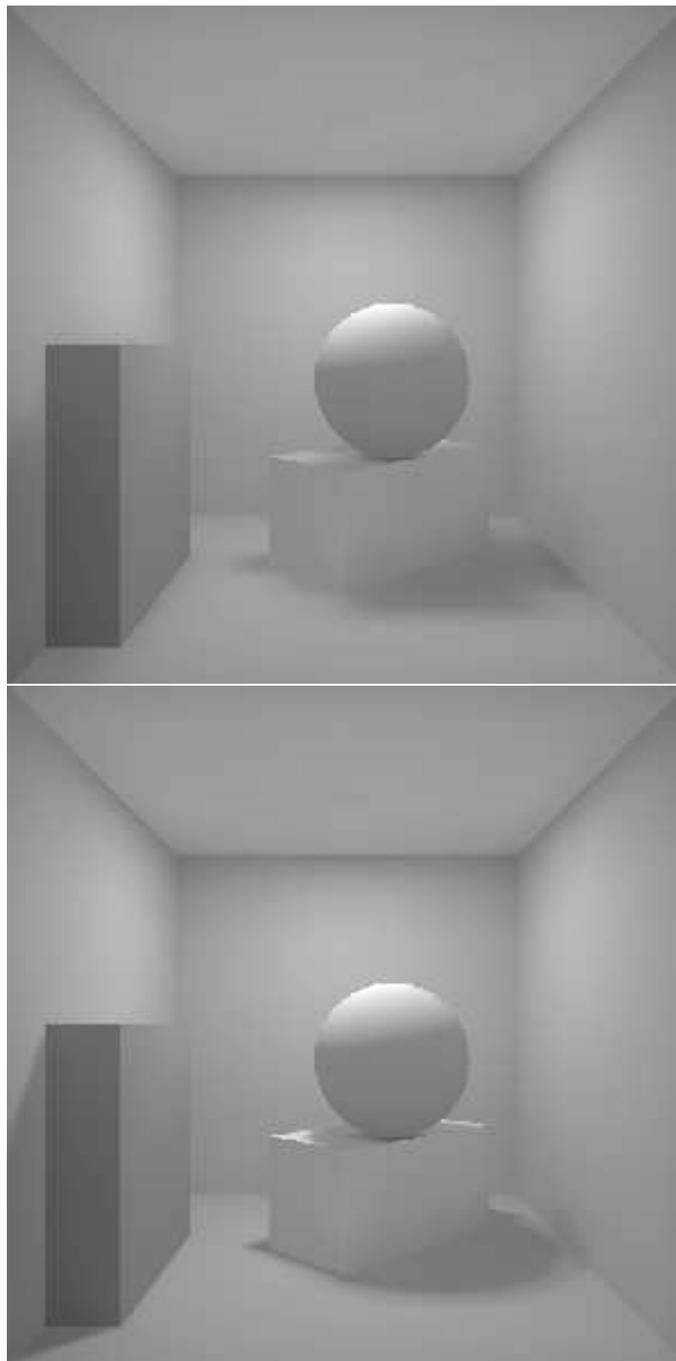


FIG. 5.4 - Images de la scène deux calculées à partir des maillages β et $\beta + 3$.

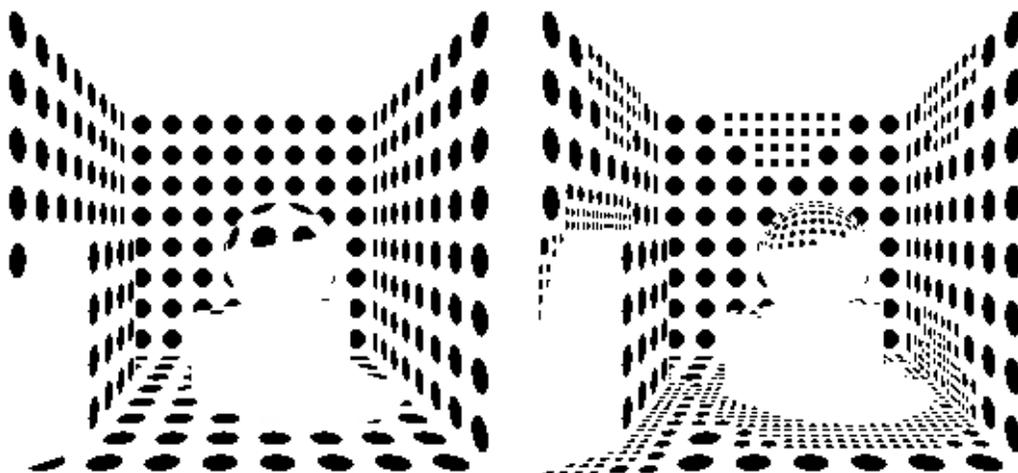


FIG. 5.5 - Deux exemples de maillage de la scène 2, l'un régulier (β à gauche) et l'autre adaptatif ($\beta + 2$ à droite). Seuls les échantillons ayant reçu un éclairage direct de la source sont tracés.

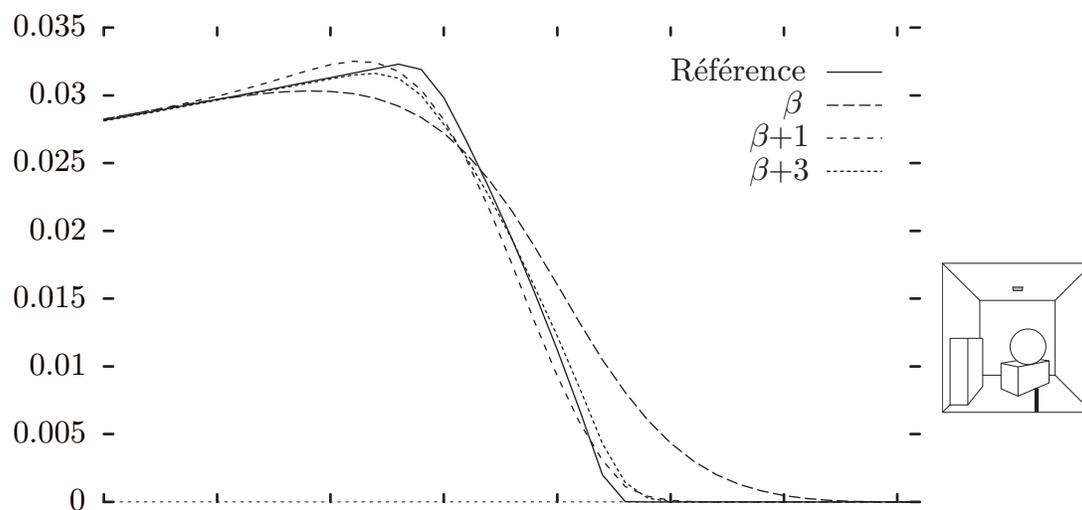


FIG. 5.6 - Profil d'éclairage dans la traversée d'une pénombre. Maillage régulier et maillages adaptatifs de profondeurs maximales croissantes.

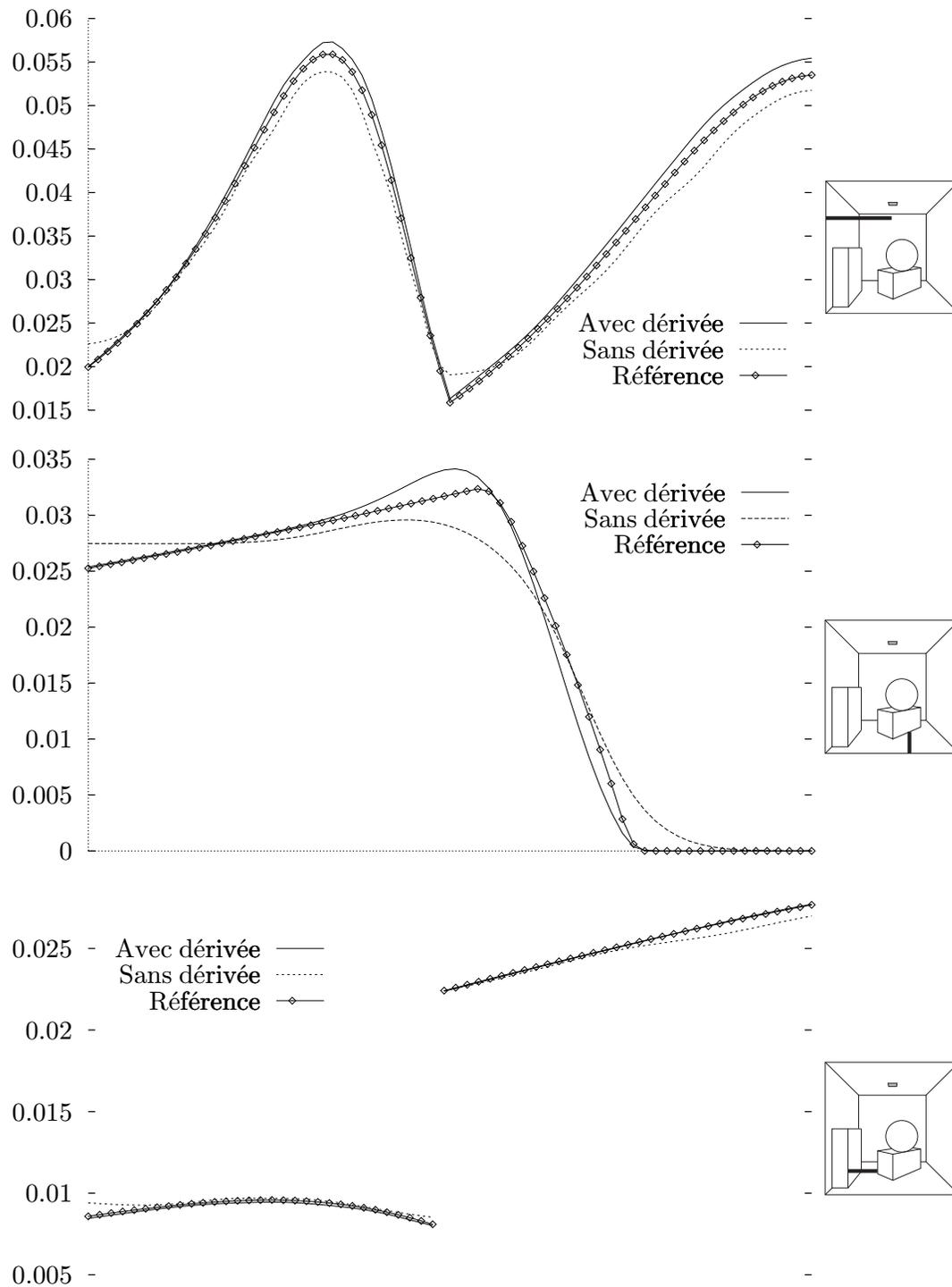


FIG. 5.7 - Utilisation des dérivées pour la reconstruction de l'éclairage direct. En haut, profil d'éclairage le long de la ligne *e*, moitié gauche. Au milieu, profil d'éclairage sur la ligne *b* à la traversée de la pénombre de la sphère. En bas, profil d'éclairage sur la ligne *c* face du parallélépipède de gauche et sol.

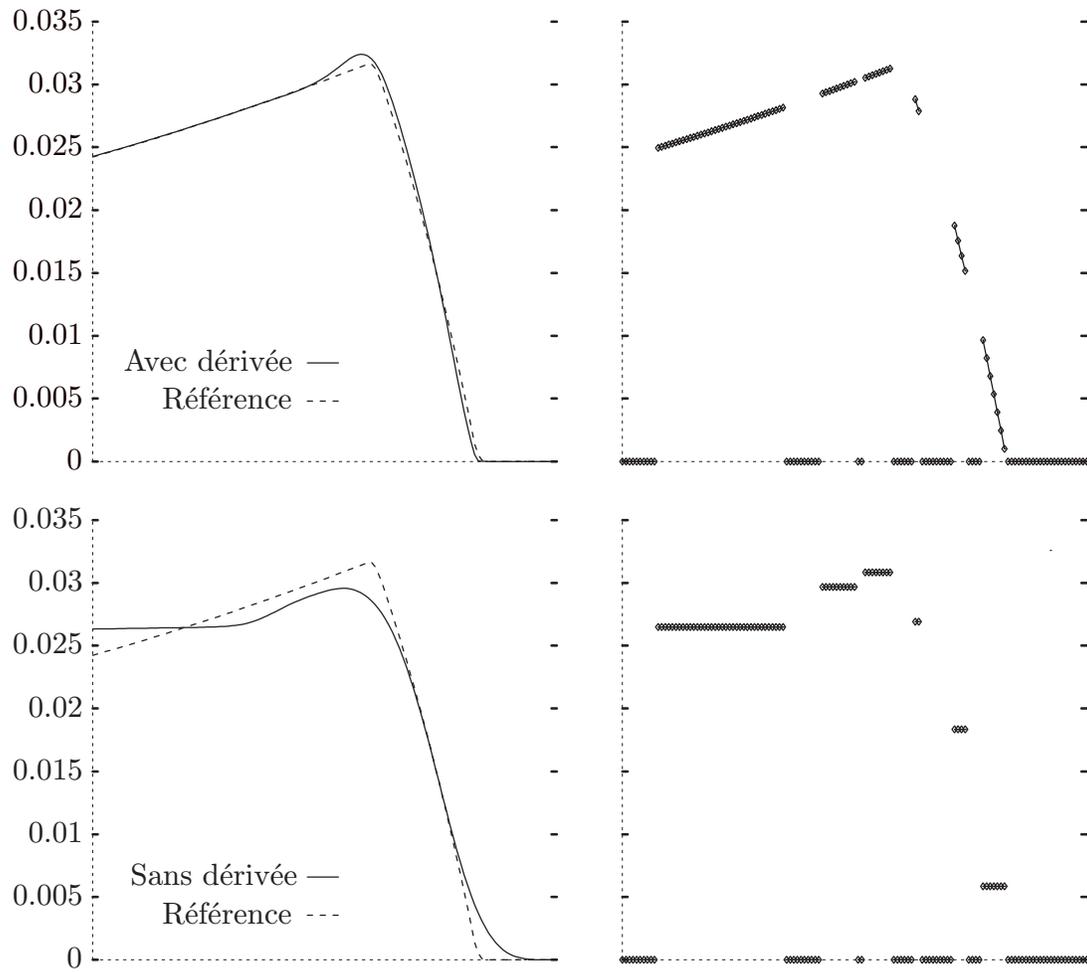


FIG. 5.8 - *Comparaison entre les radiosités reconstruites avec et sans dérivées dans une pénombre pour un même maillage de départ ($\beta + 2$). Pour les images de droites, le support des fonctions de poids a été tronqué pour séparer les différents échantillons et montrer l'extrapolation de la valeur des échantillons obtenue grâce aux dérivées.*

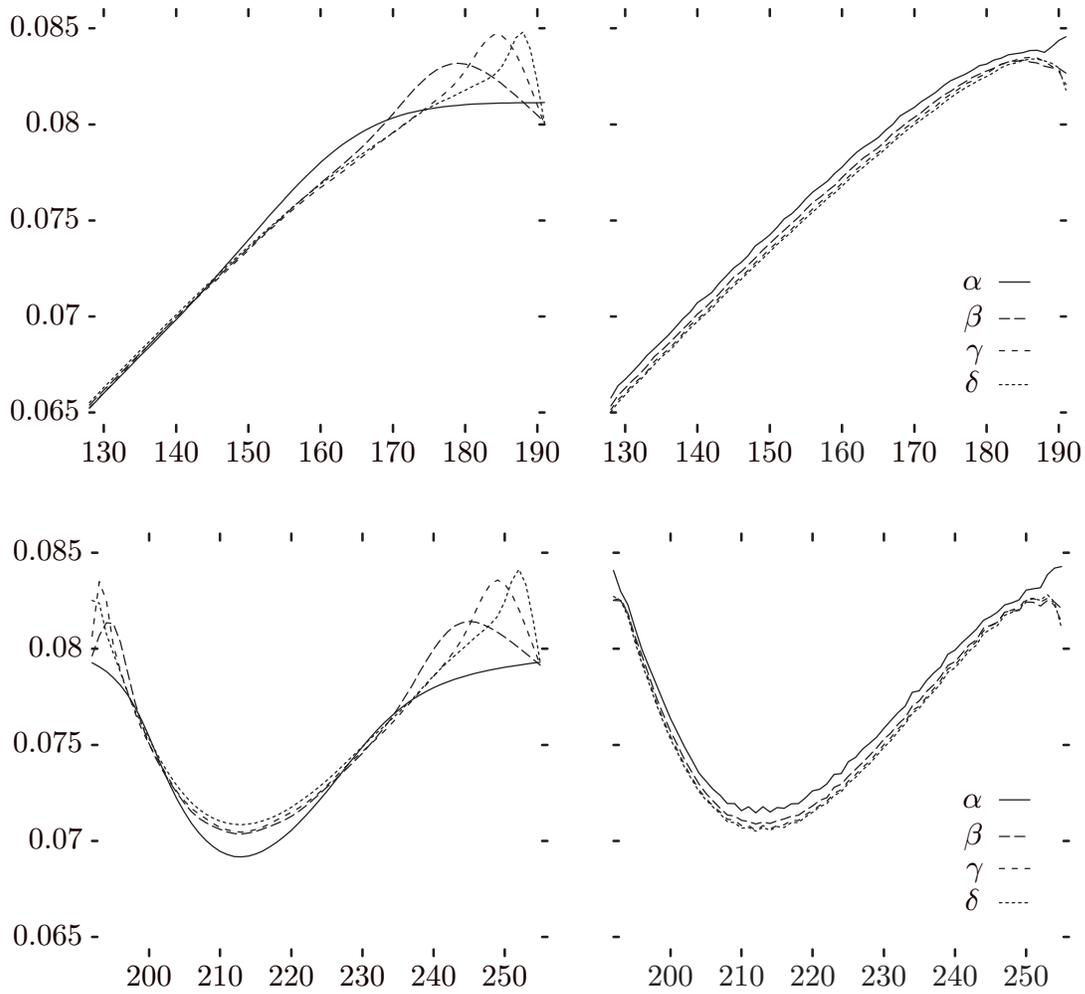


FIG. 5.9 - Profils d'éclairage indirect pour la «boite de Cornell» le long d'une ligne horizontale à mi hauteur (composante rouge). Les courbes de gauche représentent la valeur reconstruite à partir des maillages α , β , γ et δ , celles de droite, la valeur intégrée par lancer de rayon avec les mêmes maillages. Les courbes du haut représentent la moitié droite du mur du fond, celle du bas, le mur de droite. Le mur de gauche n'a pas d'éclairage dans le rouge et dans la partie gauche du mur du fond, les profils sont tous pratiquement confondus

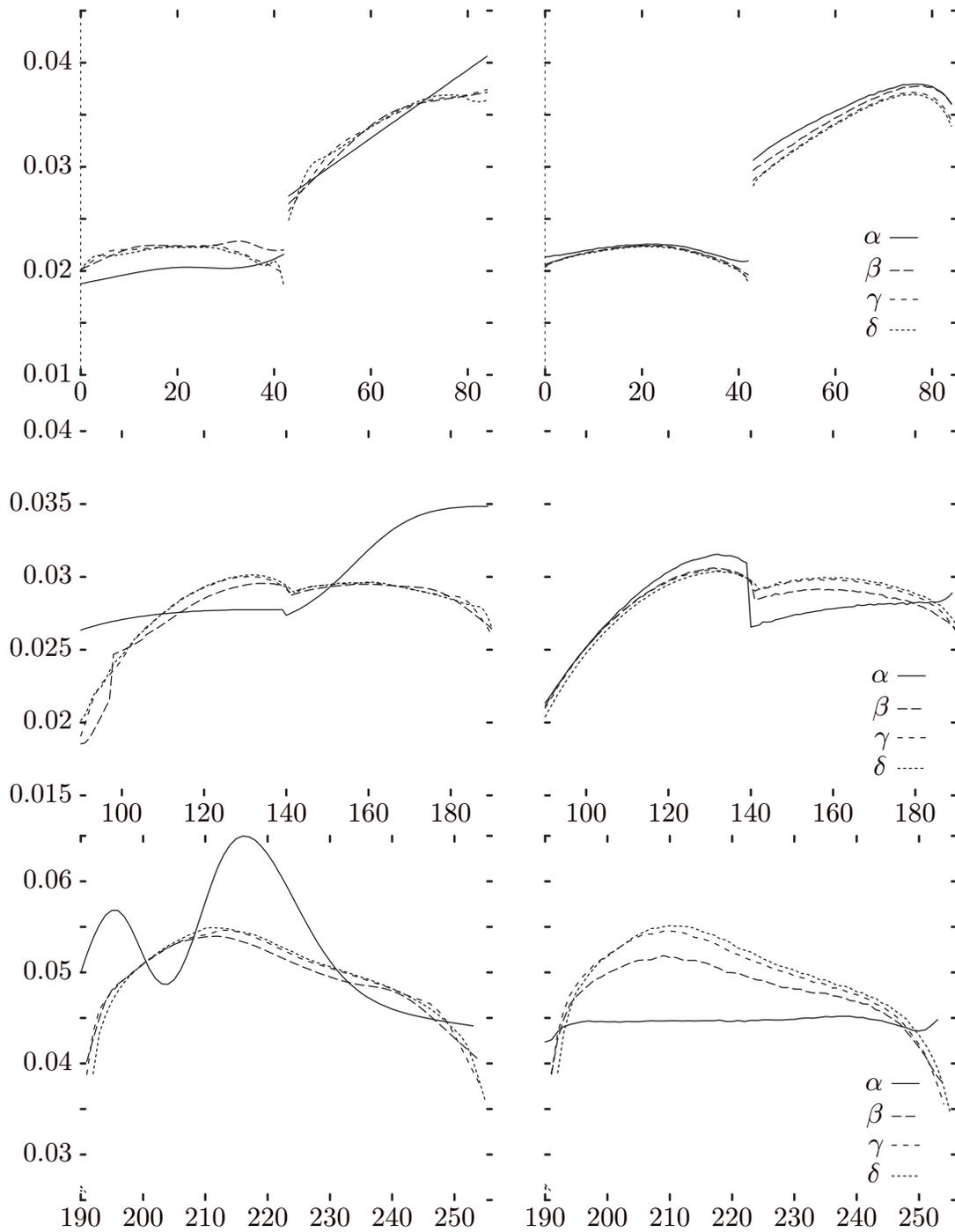


FIG. 5.10 - Profils d'éclairage indirect pour la seconde scène le long de la ligne b . La présentation est la même que pour la figure précédente. Les courbes du haut correspondent au sol et à la face avant du parallélépipède, celles du milieu à la sphère et au mur du fond et celles du bas au plafond.

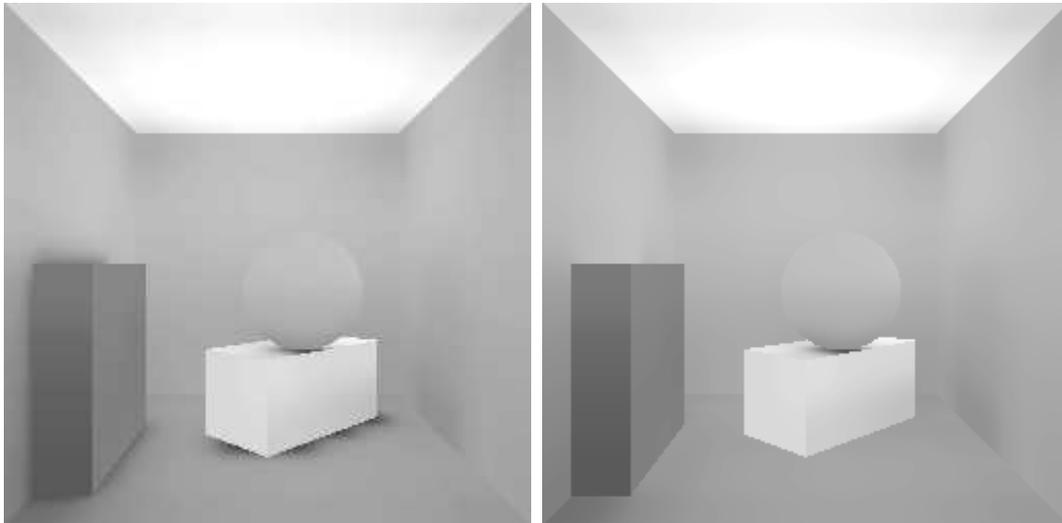


FIG. 5.11 - Élimination des fuites d'ombre. Dans l'image de gauche, tous les échantillons sont utilisés lors de la reconstruction. Dans celle de droite, les échantillons de radiosité nulle sont ignorés. Pour les deux images, l'éclairage direct a été calculé à chaque pixel par intégration de contour.

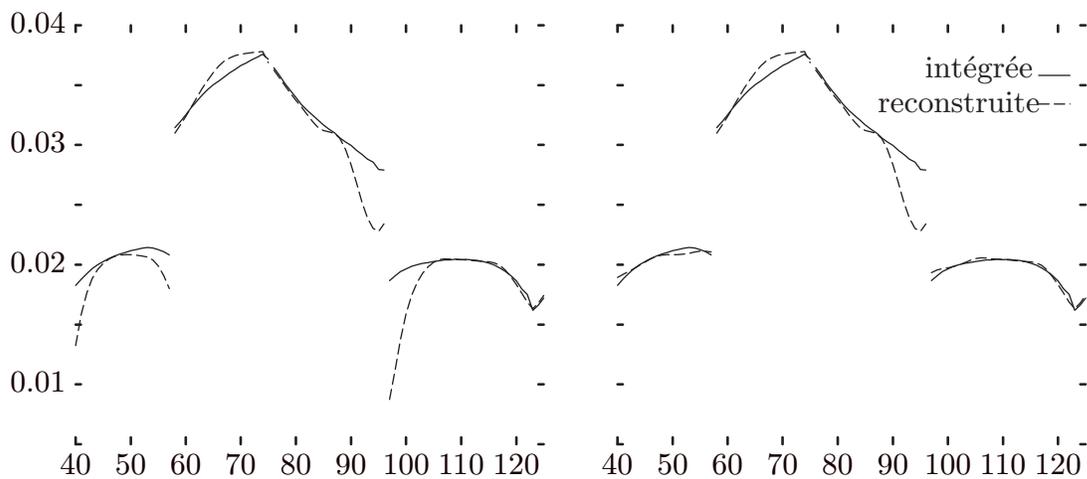


FIG. 5.12 - Profils d'éclairage le long de la ligne c pour la seconde scène. A droite, tous les échantillons sont utilisés lors de la reconstruction. A gauche on ignore les échantillons de radiosité nulle pour éliminer les fuites d'ombres.

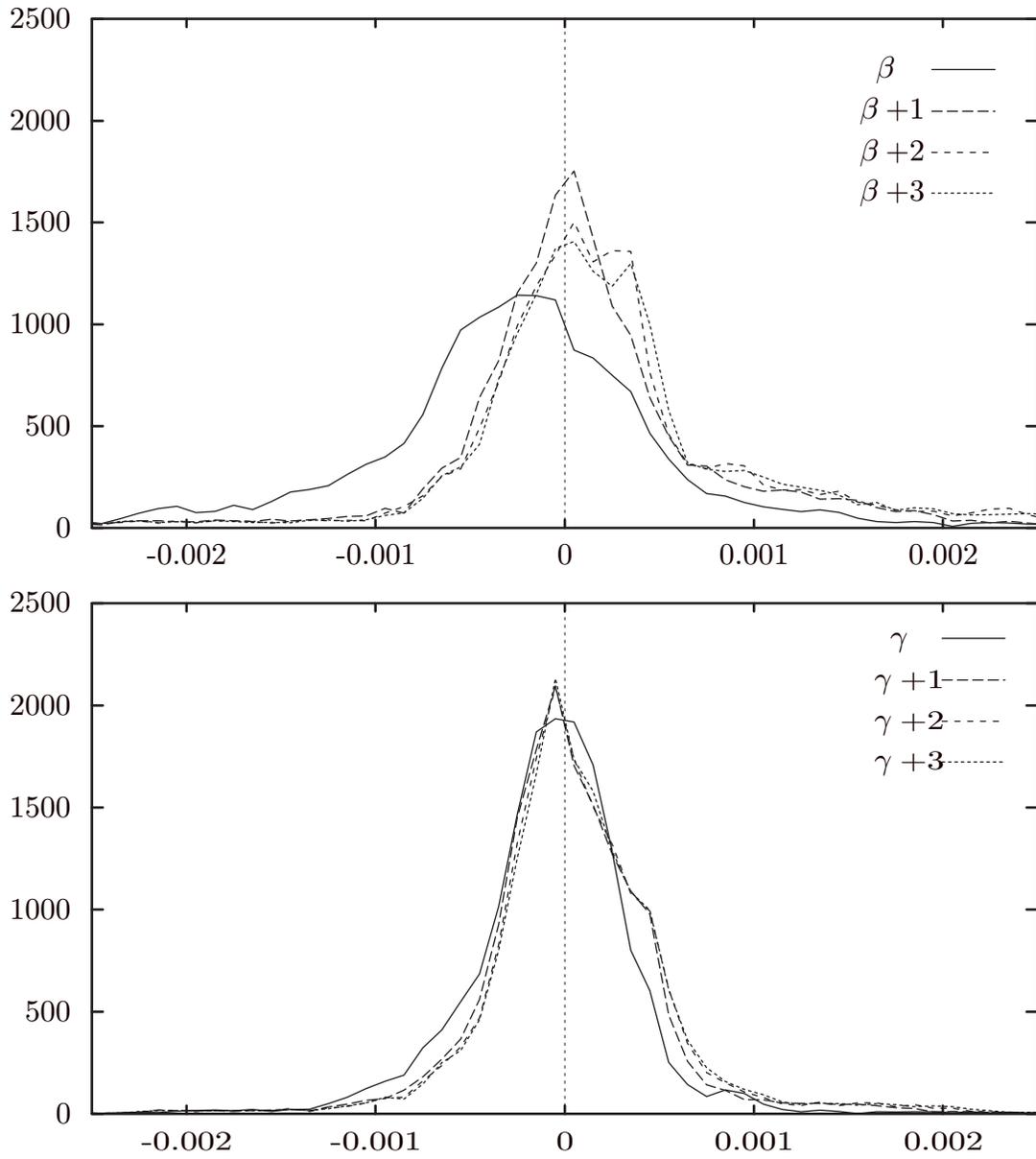


FIG. 5.13 - Histogrammes de l'écart entre valeur reconstruite et valeur intégrée pour les pixels de l'image 128×128 .

Chapitre 6

Intégration de contour approché

Nous avons vu dans le chapitre précédent qu'il est intéressant de séparer la solution de l'illumination globale d'un environnement qui fournit l'éclairage indirect et le calcul de l'éclairage direct. Dans ce chapitre, nous présentons une méthode précise d'intégration du facteur de forme destinée à être utilisée, dans un lancer de rayon, pour le calcul pixel par pixel de l'ombrage d'une source étendue. La notion de base est celle de contour approché qui permet d'étendre simplement l'utilisation de l'intégration de contour à des environnements non polyédriques.

Dans la section 6.1 sont d'abord présentées les solutions déjà proposées au problème de l'éclairage par des sources étendues ainsi que les motivations pour l'emploi de l'intégration de contour dans le cas du calcul pixel par pixel. La section 6.2 décrit la construction du contour approché qui est réalisée en adaptant un algorithme de *marching cubes* destiné à l'origine à la construction d'approximations polyédriques de surfaces isopotentielles. La technique utilisée pour vérifier la validité de l'approximation faite et raffiner éventuellement le contour est l'objet de la section 6.3. Les sections 6.4 et 6.5 présentent quelques points d'implémentation et l'extension de la méthode au calcul des dérivées. Enfin, nous comparons les résultats obtenus à ceux d'un algorithme de Monte Carlo adaptatif.

6.1 Motivations

Cette section reprend les algorithmes décrits par Woo *et al.* dans [WPF90]. Cet état de l'art des algorithmes d'ombrage distingue les méthodes suivantes :

1. Algorithme de frame buffer : l'algorithme de Brotman et Badler [BB84] distribue de manière stochastique plusieurs sources ponctuelles sur la surface de la source étendue. Pour chacune de ces sources, l'algorithme des volumes d'ombre de Crow [Cro77] est appliqué et les différentes contributions sont ensuite accumulées. Le principal inconvénient de cette approche est que le choix des points et de leur emplacement est le même pour tous les pixels de l'image et n'est donc pas partout bien adapté.

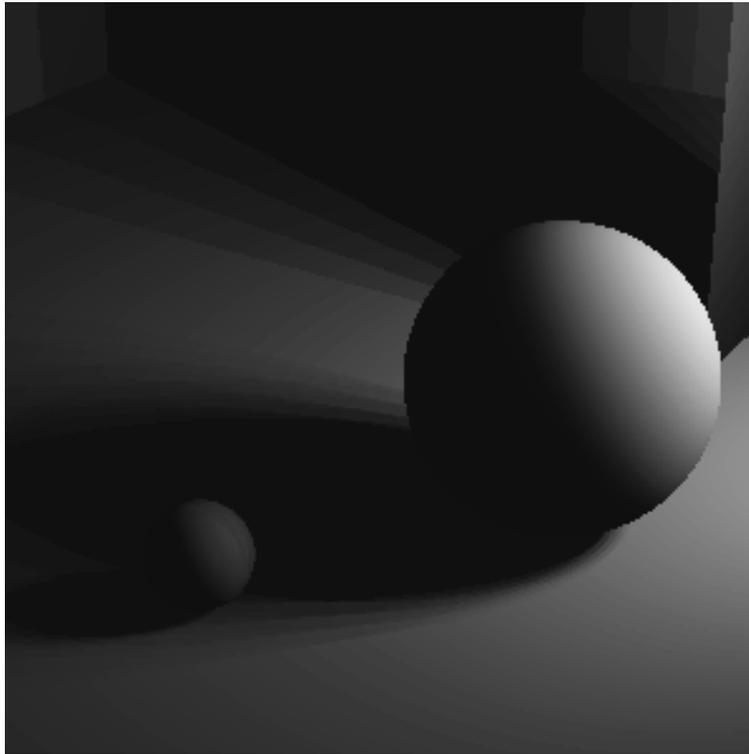


FIG. 6.1 - *Effet de marches d'escalier dû à l'approximation de la source étendue par un nombre trop faible de sources ponctuelles (16 dans ce cas).*

2. Lancer de rayon distribué : cette méthode proposée par Cook *et al.* [CPC84] lance plusieurs rayons vers des positions perturbées de manière aléatoire sur la source. Plusieurs méthodes de choix des rayons nécessaires (nombre et position) ont été proposées : estimation de la variance [DW85], échantillonnage stratifié [LRU85]. Récemment, Kok et Jansen [KJ92] ont proposé un algorithme adaptatif faisant une synthèse de ces approches. Ces méthodes bénéficient des caractéristiques particulières de l'oeil humain qui rendent le bruit beaucoup plus tolérable que les effets d'aliasage.
3. Tracé de cônes : en 1984, Amanatides propose de remplacer la droite utilisée en lancer de rayons par un cône. Cette technique permet de traiter des lampes sphériques en lançant un cône d'ombrage. Quand un seul obstacle bloque la source, les pénombres obtenues sont de bonne qualité car la fraction de source cachée est calculée précisément par la routine d'intersection entre cône et objet. Par contre, la méthode ne garantit pas un résultat correct si plusieurs obstacles sont présents simultanément.
4. Intégration de la surface visible : pour des scènes exclusivement composées de polygones, Woo *et al.* proposent d'utiliser un algorithme de visibilité

exacte (clipping) de type algorithme d'Atherton [AWG78] pour calculer la partie visible de la source. Cet algorithme peut être rendu plus efficace par la construction et l'emploi de listes restreintes d'objets candidats.

5. algorithmes de radiosité. L'éclairage est calculé sur un maillage puis interpolé. Les ombres franches nécessitent un maillage très fin pour éviter les artefacts dans l'image.

Les algorithmes d'échantillonnage aléatoire permettent de modéliser de nombreux effets et fournissent de bons résultats dans le calcul des pénombres de sources étendues. Cependant, pour ce problème particulier, la fonction à échantillonner varie régulièrement en dehors des discontinuités et il semble possible d'exploiter ces faibles variations pour accélérer encore les calculs.

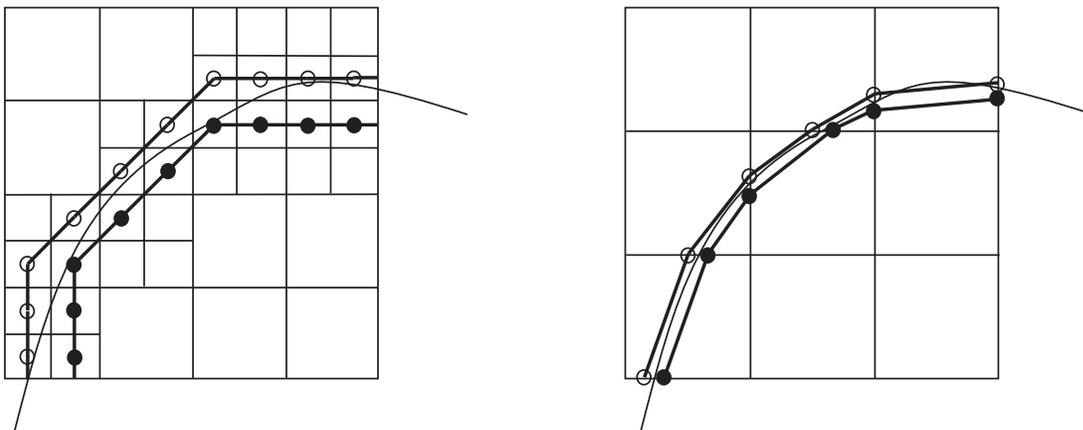


FIG. 6.2 - *Comparaison entre subdivision adaptative de la surface et approximation de son contour. Avec le même nombre de rayons, la précision sur la position du contour sur la source est meilleure.*

Quand le facteur de forme est calculé par lancer de rayons, la source est remplacée par plusieurs lampes ponctuelles. Si ces lampes sont trop peu nombreuses, l'éclairage dans les pénombres prend la forme de marches d'escalier (voir figure 6.1). L'échantillonnage aléatoire remplace cet effet par du bruit mais comme noté dans [KJ92], il peut aussi être nécessaire de raffiner la frontière entre partie visible et partie invisible sur la source. Si cette frontière est assez lisse, il est alors plus intéressant de l'approcher par un contour polygonal plutôt que par un contour «en marche d'escalier» (voir figure 6.2), c'est-à-dire d'utiliser une intégration par la méthode des trapèzes plutôt que par celle des rectangles. De plus, en utilisant le même nombre de rayons, la position du contour peut être connue avec plus de précision ce qui permet, dans la pénombre, de rendre distinctes les valeurs de deux pixels voisins. Nous présentons maintenant la technique utilisée pour construire cette approximation polygonale du contour.

6.2 Construction du contour approché

6.2.1 Principe

Dans les zones d'où l'on voit toute la source, le contour est simplement le bord de la source. Dans les ombres, il est réduit à rien. Les zones intéressantes pour le calcul sont donc les pénombres. Le contour visible d'une source \mathcal{S} se compose alors en général de morceaux de silhouette des obstacles se projetant sur l'intérieur de \mathcal{S} et d'une partie du bord de \mathcal{S} . Pour simplifier l'algorithme, ces deux types de contribution sont calculées séparément. Nous présentons d'abord le cas de la partie intérieure à \mathcal{S} .

Le contour polygonal approché est construit par une variante de l'algorithme des *marching cubes* en dimension deux que nous appelons donc *marching squares*. On s'intéresse ici à une source rectangulaire, le traitement des sources polygonales plus compliquées étant évoqué plus tard. Au départ, la source est subdivisée régulièrement en petits rectangles. Les sommets de ces rectangles $S_{j,k}$, $j = 0, \dots, p$ et $k = 0, \dots, q$ sont organisés en grille. On définit la visibilité vis_M comme la fonction caractéristique de la partie visible de la source depuis le point M où l'on cherche l'illumination et on la calcule en chaque sommet par lancer de rayon. Ces rayons lancés au départ sont appelés rayons primaires dans la suite. Si vis_M prend des valeurs différentes en deux points, le contour traverse forcément le segment joignant ces points. On va utiliser cette propriété pour détecter le passage du contour dans les rectangles $(S_{j,k}, S_{j+1,k}, S_{j+1,k+1}, S_{j,k+1})$.

6.2.2 Marching squares

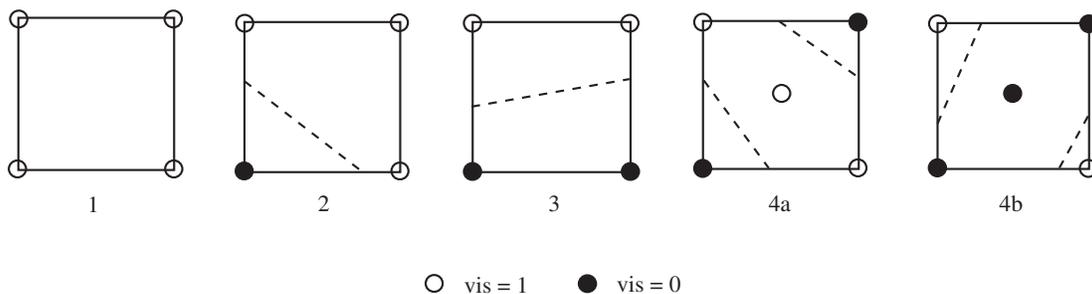


FIG. 6.3 - Configurations de passage du contour.

Comme chaque sommet est soit vu, soit caché, il y a *a priori* 16 configurations possibles pour les quatre sommets d'un rectangle. Par symétries, on réduit ce nombre aux quatre cas distincts suivants (voir figure 6.3) :

1. les quatre sommets sont simultanément vus ou cachés. Aucun contour n'est détecté.

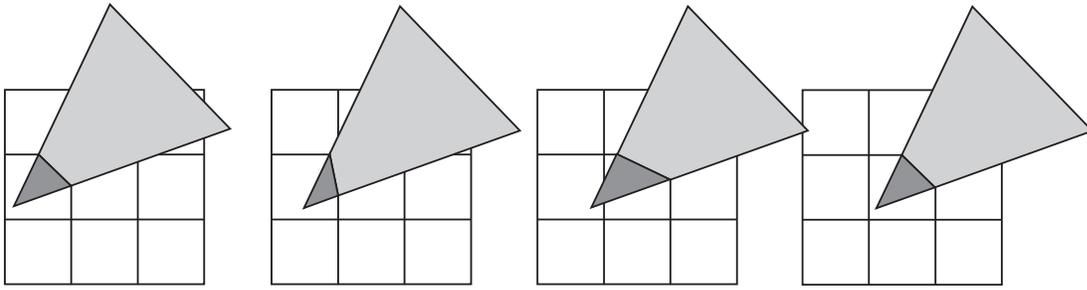


FIG. 6.4 - Variation périodique de la surface de source négligée.

2. un des sommets est différent des trois autres. Le contour traverse deux arêtes adjacentes du rectangle.
3. deux sommets voisins sont vus et les deux autres non. Le contour traverse deux arêtes opposées du rectangle.
4. deux sommets diamétralement opposés sont vus et les deux autres non. Il y a deux possibilité et le choix est fait en lançant un rayon supplémentaire au centre du rectangle.

Dans chacun des cas où cela est nécessaire, on construit un segment (cas 2 et 3) ou deux (cas 4) du contour polygonal approché. La procédure utilisée pour localiser les extrémités du segment est la suivante. Elle consiste à réduire la taille du segment par lequel on sait que le contour passe en utilisant la dichotomie. Par exemple, si $vis_M(S_{j,k}) = 1$ alors que $vis_M(S_{j+1,k}) = 0$, on lance un rayon vers le milieu de ces deux points. Suivant sa visibilité, on va recommencer l'opération sur l'un des deux demi-segments puis récursivement jusqu'à ce que la longueur de l'intervalle soit plus petite qu'un seuil donné. En pratique, comme les rectangles ont la même taille au départ, ce seuil correspond à un nombre d'étapes de dichotomie fixé. Le milieu du dernier segment devient un sommet du contour polygonal approché. Pour chaque rectangle, on obtient ainsi un segment $[C_i, C_{i+1}]$ du contour polygonal (ou deux) dont la contribution au vecteur de forme total est donnée par $\beta_i \vec{n}_i / 2\pi$ où β_i est l'angle (C_i, M, C_{i+1}) et \vec{n}_i est la normale unitaire au plan défini par M, C_i et C_{i+1} .

6.3 Raffinement du contour approché

Si la vraie silhouette de l'obstacle et donc le vrai contour n'ondule pas trop, on obtient une bonne approximation du vecteur de forme. Cependant, si cette silhouette présente, par exemple, un angle dont la pointe se projette à l'intérieur d'un des rectangles, cette pointe sera coupée du contour. L'erreur ainsi introduite provoque en plus un aliassage important. Quand on traverse la pénombre, la projection de la pointe glisse de rectangle en rectangle et la quantité coupée

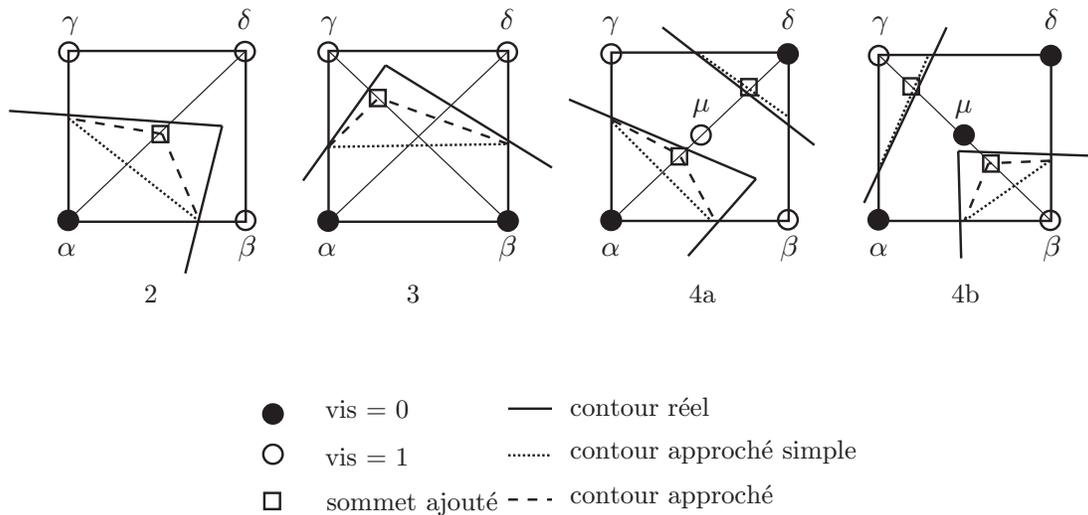


FIG. 6.5 - *Sommets ajoutés au contour pour tester sa validité.*

oscille périodiquement (voir figure 6.4). Ces cas doivent donc être détectés et le contour polygonal ajusté en conséquence. Pour cela, l'algorithme découpe chacun des segments obtenus dans la première phase en deux sous-segments. Le point intermédiaire rajouté est aussi calculé par dichotomie indépendamment des extrémités du segment. La variation de facteur de forme quand on passe du contour en un segment au contour en deux segments sert à décider si un raffinement plus important est nécessaire.

6.3.1 Critère

Pour ajouter un sommet intermédiaire, on utilise encore la dichotomie. Il faut donc choisir un couple de points de visibilité différentes, et réduire le segment associé. Ce choix dépend du cas (2, 3 ou 4) dans lequel on était précédemment (voir figure 6.5).

Cas 2 un sommet est calculé sur la diagonale du rectangle (dichotomie en partant de $[\alpha, \delta]$);

Cas 3 un sommet est calculé sur l'une des deux diagonales du rectangle choisie au hasard (dichotomie en partant de $[\alpha, \delta]$ ou $[\beta, \gamma]$);

Cas 4 un sommet est calculé sur chacune des deux demi-diagonales du rectangle. La valeur de $vis_M(\mu)$ est déjà connue (dichotomies en partant de $[\alpha, \mu]$ et $[\mu, \delta]$ ou $[\beta, \mu]$ et $[\mu, \gamma]$ suivant le sous cas).

Pour un rectangle donné, si on fait la différence entre les contributions du segment seul et des deux nouveaux segments, on obtient le vecteur de forme du

petit triangle compris entre les deux «versions» du contour. Soit δF le facteur de forme correspondant. Pour un vrai contour rectiligne, δF est presque nul et il augmente en général quand le rayon de courbure du contour diminue. On va donc comparer δF à un seuil pour décider si il faut raffiner ou non le contour dans le rectangle.

Soit ϵ la longueur des petits intervalles obtenus en fin de dichotomie. Comme les extrémités du segment ne sont connus qu'à ϵ près, un vrai contour rectiligne ne sera quoi qu'il arrive calculé qu'approximativement. L'incertitude sur l'intégration correspond au facteur de forme d'une bande de largeur ϵ autour du segment. C'est cette valeur qui est estimée en utilisant la formule du disque (§1.2.1) et sert de seuil, ce qui permet un ajustement automatique en fonction du nombre de pas de dichotomie employé.

Pour affiner la contribution du rectangle, on le découpe en quatre et les sous-rectangles sont ensuite traités tour à tour par la même méthode. Le nombre de pas de dichotomie est décrémenté de un ce qui laisse ϵ constant. Comme la longueur de contour traversant le sous-rectangle est deux fois plus courte, le seuil est divisé par deux. Cette opération peut être appliquée récursivement si nécessaire.

6.3.2 Problèmes avec le raffinement

Quand un rectangle est découpé, des sommets sont ajoutés au milieu de chacune de ses arêtes. Si les sommets de l'arête ont la même visibilité, il est possible que celle du sommet ajouté soit différente. Par exemple, sur la partie gauche de la figure 6.6, le rectangle de gauche R_1 sera ou a déjà été traité et ignoré puisque ces quatre sommets ont la même visibilité. Lors du traitement du rectangle de droite R_2 , on décide de le subdiviser et le nouveau sommet S ajouté sur l'arête centrale montre que la situation de R_1 est plus compliquée (le contour traverse l'arête deux fois et non pas zéro). R_1 doit donc lui aussi être subdivisé. Le cas de la figure 6.6 à gauche est simple car même si le traitement de R_1 a déjà été effectué, il n'y a pas eu de contribution au facteur total. Par contre, dans la partie droite, la contribution de R_1 correspondant au segment C doit être soustraite pour permettre la subdivision et le recalcul sur les sous-rectangles. La découverte d'une incohérence peut aussi se faire avec des différences de niveau plus importante dans l'arbre et alors que le rectangle voisin n'existe pas à cette profondeur (voir le bas de la figure 6.6).

Dans [Ved93], nous avons présenté une implémentation simplifiée pour laquelle la découverte d'une incohérence déclenche la reprise à zéro de tout le calcul du vecteur de forme à partir d'une grille initiale de rectangles plus fine. Nous décrivons ici une procédure moins radicale.

La source est en fait représentée par une forêt d'arbres quaternaires organisés suivant une grille et dont les noeuds sont les rectangles. On peut donc étendre

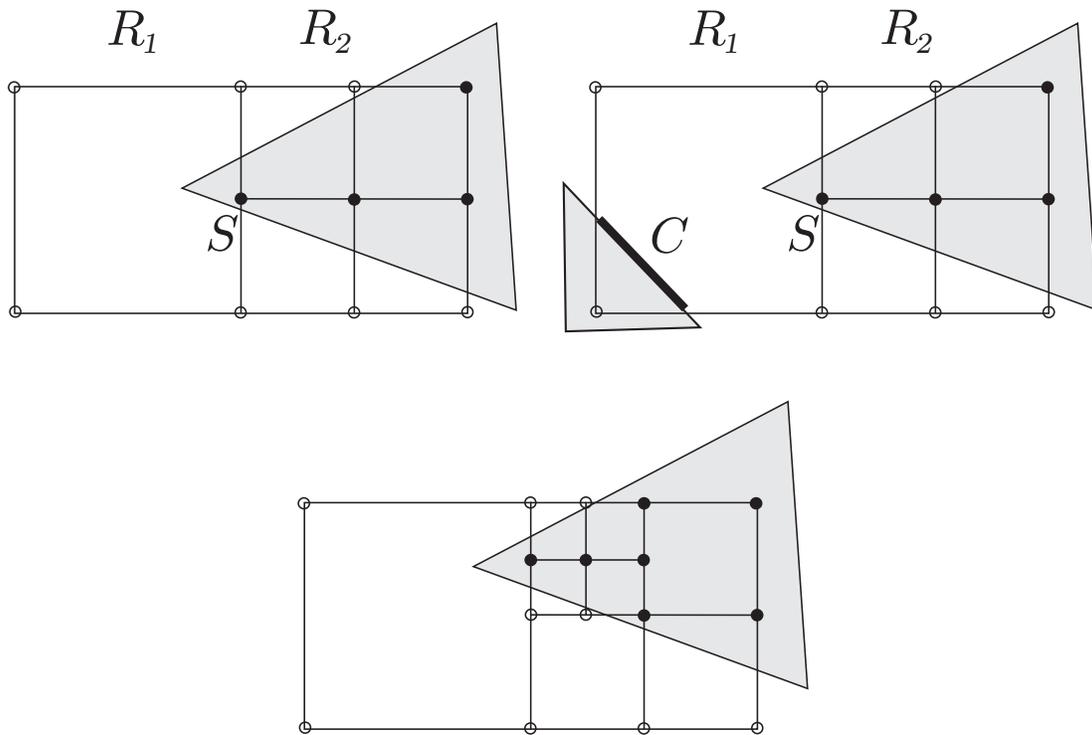


FIG. 6.6 - *Incohérences créées par le raffinement d'un rectangle.*

la notion de voisinage naturellement et en pratique considérer l'ensemble comme un seul arbre. La première mesure à prendre consiste à éliminer les cas vraiment difficiles, c'est-à-dire pour lesquels on trouve une incohérence entre rectangles de niveaux très différents. Pour cela, on restreint l'arbre quaternaire. Pour imposer cette propriété, il suffit avant de subdiviser une feuille, de vérifier la profondeur de ses voisins et de subdiviser aussi celles qui sont trop hautes.

L'algorithme de raffinement utilise une pile des rectangles à traiter sur laquelle sont placés au début tous ceux du niveau 0. On les dépile ensuite un par un pour les traiter. L'algorithme termine quand la pile est vide. Avant de traiter un rectangle, on vérifie qu'il ne l'a pas déjà été. En effet, si une incohérence a été détectée, le rectangle voisin peut avoir été subdivisé et ses sous-rectangles traités plus tôt. Le traitement d'un rectangle comprend 3 opérations :

1. Calculer les deux «versions» du contour ;
2. Calculer δF ;
3. Si δF est plus grand que le seuil, subdiviser le rectangle, sinon stocker la contribution du contour dans le noeud.

La subdivision d'un rectangle comprend elle 4 opérations :

1. Si l'un des voisins est trop haut dans l'arbre, le subdiviser ;

2. Calculer, si nécessaire, la visibilité des nouveaux sommets ;
3. Pour chaque arête où une incohérence est détectée, si le voisin est une feuille, le marquer comme traité et le subdiviser ;
4. Empiler les quatre sous-rectangles.

La première opération assure que l'arbre quaternaire reste restreint. Pour éviter d'avoir à soustraire les contributions des rectangles déjà traités que l'on subdivise, les vecteurs de forme ne sont pas tout de suite accumulés mais simplement stockés dans les noeuds de l'arbre. Quand l'algorithme de raffinement a terminé, on parcourt les feuilles de l'arbre pour rassembler le vecteur total.

6.4 Remarques sur l'implémentation

6.4.1 Structures de données

La forêt d'arbres quaternaires complète est construite en mémoire au départ. On garde aussi dans un tableau tous les sommets des noeuds de l'arbre, qui sont donc naturellement partagés entre rectangles voisins. Quand on subdivise un rectangle, certains des nouveaux sommets ont déjà été utilisés dans les opérations de dichotomie et leur visibilité est disponible sans calcul. Si le nombre d'étapes de dichotomie est élevé, la taille de ce tableau peut poser problème. On peut alors ne garder que les premiers niveaux qui sont les seuls utiles en cas de redémarrage du calcul.

6.4.2 Alternative à la dichotomie

Dans la méthode décrite, la dichotomie sert à préciser la position du contour de l'objet en projection sur la source. Elle a l'avantage d'utiliser le lancer de rayon comme seule primitive de test de visibilité, ce qui impose peu de contraintes sur les objets accessibles à la méthode. Cependant, pour beaucoup d'objets, il est possible de déterminer exactement et plus efficacement le point de passage du contour. Les calculs nécessaires sont différents de ceux d'intersection avec un rayon mais ne sont pas beaucoup plus compliqués. Nous présentons ici le principe de ces calculs et avec plus de détails le cas d'un polygone convexe, d'une surface implicite et d'une surface paramétrique.

Rappelons d'abord les données du problème. L'illumination est calculée au point M et on dispose de deux échantillons $S_{j,k}$ et $S_{j,k+1}$ de visibilités différentes situés sur la source. L'échantillon $S_{j,k+1}$ est, par exemple, caché par un objet O . On cherche le rayon extrême de l'ensemble des rayons issus de M , contenus dans le plan défini par M , $S_{j,k}$ et $S_{j,k+1}$ et qui intersectent O . En général, le contact de ce rayon avec O est réduit à un point. Une fois que le rayon extrême pour

l'objet O est obtenu, il faut le relancer en ignorant O pour tester l'intersection avec d'autres obstacles éventuels.

Cas d'un polygone convexe

Le plan support du triangle $(M, S_{j,k}, S_{j,k+1})$ et le plan support du polygone définissent une droite qui intersecte le bord du polygone en deux points A et B (voir la figure 6.7). Le rayon extrême recherché passe par M et un de ces deux points d'intersection. En pratique, on peut parcourir les arêtes du polygone et calculer les positions des sommets successifs par rapport au plan du triangle $(M, S_{j,k}, S_{j,k+1})$ pour trouver celles qui percent ce plan puis ordonner les deux intersections dans la direction $(S_{j,k}, S_{j,k+1})$ pour choisir la bonne.

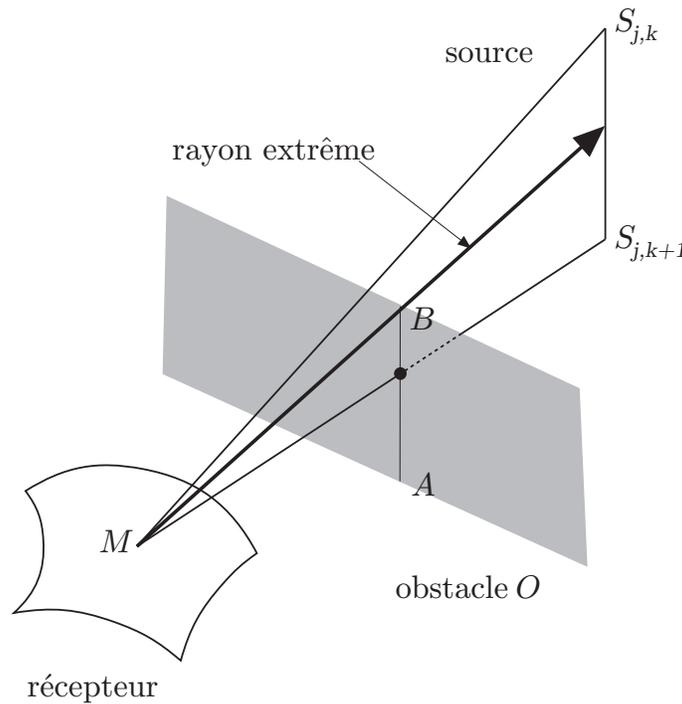


FIG. 6.7 - Rayon limite pour un polygone.

Cas d'une surface implicite

On suppose que la surface est définie par l'équation implicite $f(x, y, z) = 0$ (voir la figure 6.8). Alors la position du point de tangence du rayon extrême sur l'objet O est définie par les trois équations :

$$\begin{cases} \alpha x + \beta y + \delta z + \gamma & = 0, \\ f(x, y, z) & = 0, \\ (x - x_M)f_x(x, y, z) + (y - y_M)f_y(x, y, z) + (z - z_M)f_z(x, y, z) & = 0, \end{cases}$$

où $(\alpha, \beta, \delta, \gamma)$ définissent l'équation du plan support du triangle $(M, S_{j,k}, S_{j,k+1})$ et la troisième équation exprime la tangence du rayon. Par substitution d'une des variables à partir de la première équation, on se ramène à un système de deux équations en général non linéaires à deux inconnues. Ce système peut être résolu en utilisant une méthode itérative de type Newton-Raphson, le point d'intersection du rayon $(M, S_{j,k+1})$ avec O déjà connu fournissant un bon point de départ à l'algorithme. Comme dans le cas du lancer de rayon, des méthodes plus simples peuvent être mises au point dans des cas particuliers comme la sphère.

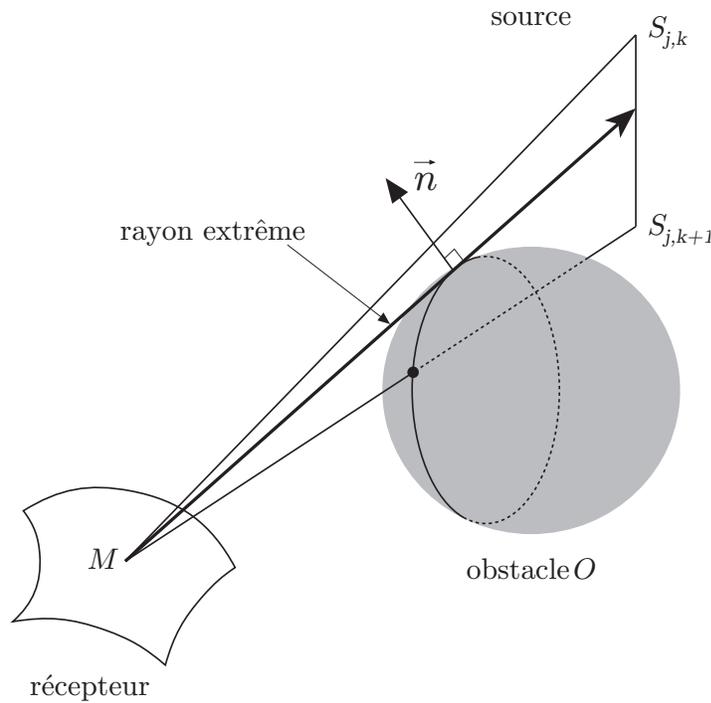


FIG. 6.8 - Rayon limite pour une sphère.

Cas d'une surface paramétrique

On suppose que la surface est définie par la fonction \vec{F} :

$$\begin{aligned} [a, b] \times [c, d] &\rightarrow \mathbb{R}^3 \\ (u, v) &\rightarrow \vec{F}(u, v). \end{aligned}$$

Alors la position du point de tangence du rayon extrême sur l'objet O est définie par les deux équations suivantes qui expriment respectivement la tangence du rayon à la surface et son appartenance au plan $(M, S_{j,k}, S_{j,k+1})$:

$$\begin{cases} \left(\frac{\partial \vec{F}}{\partial u} \times \frac{\partial \vec{F}}{\partial v} \right) \cdot \overrightarrow{MF} = 0 \\ \begin{pmatrix} \alpha \\ \beta \\ \delta \end{pmatrix} \cdot \overrightarrow{MF} = 0, \end{cases}$$

et on est ramené comme précédemment à un système de deux équations à deux inconnues. Si l'objet est constitué de plusieurs carreaux de surface assemblés, la résolution itérative du système doit être arrêtée aux passages de carreau à carreau. Si la normale est discontinue au travers de la frontière, il faut vérifier que le point de contour ne se trouve pas sur cette frontière avant de continuer la résolution sur le carreau suivant.

Une fois que le rayon extrême pour l'objet O est calculé, il faut le relancer en ignorant O pour vérifier qu'il n'y a pas d'autre obstacle et éventuellement recommencer la procédure de calcul du rayon extrême. On dispose à la fin de la position exacte du contour projeté sur la source et sa distance à la source ce qui améliore la précision du calcul des dérivées décrit maintenant.

6.5 Calcul des dérivées

La méthode présentée dans ce chapitre a été prévue pour calculer l'éclairage direct d'une source étendue pixel par pixel. Elle peut aussi être utilisée pour calculer les vecteurs de forme avec les sources primaires dans une radiosité progressive. Nous donnons donc les formules de calcul des dérivées pour permettre l'utilisation de la méthode de maillage décrite au chapitre 4.

On rappelle que la contribution d'un segment $[C_i, C_{i+1}]$ de contour est

$$\vec{G}_i = \beta_i \vec{n}_i, \quad (6.1)$$

Si les coordonnées de M sont (x, y, z) dans le système $(\vec{I}, \vec{J}, \vec{K})$, la dérivée partielle de \vec{G}_i par rapport à x s'écrit

$$\partial_x \vec{G}_i = \partial_x \beta_i \vec{n}_i + \beta_i \partial_x \vec{n}_i, \quad (6.2)$$

avec

$$\partial_x \beta_i = \frac{\vec{c}_i + c_{i+1} \vec{c}_i - (\vec{c}_i \cdot c_{i+1} \vec{c}_i) \left(\frac{\vec{c}_i}{c_i^2} + \frac{c_{i+1} \vec{c}_i}{c_{i+1}^2} \right)}{\sqrt{\sqrt{c_i^2} \sqrt{c_{i+1}^2} - \vec{c}_i \cdot c_{i+1} \vec{c}_i}} \cdot \vec{I}, \quad (6.3)$$

et

$$\partial_x \vec{n}_i = \frac{\partial_x \vec{w}_i - (\vec{n}_i \cdot \partial_x \vec{w}_i) \vec{n}_i}{\|\vec{w}_i\|}, \quad (6.4)$$

où on a posé $\vec{c}_i = \vec{M}C_i$, $\vec{w}_i = \vec{c}_i \times c_{i+1} \vec{c}_i$ et avec $\partial_x \vec{c}_i = C_i \vec{C}_{i+1} \times \vec{I}$.

6.6 Résultats

Nous présentons maintenant les résultats obtenus avec la méthode décrite en les comparant à ceux d'un algorithme de Monte Carlo adaptatif. L'algorithme a été utilisé dans le calcul, par lancer de rayons, d'images d'une scène comprenant une source rectangulaire. Dans la section 6.6.1, nous rappelons le principe de l'algorithme de Kok et Jansen qui a été implémenté pour comparaison. La section 6.6.2 décrit le traitement d'une ligne de pixels de l'image ainsi que les optimisations qui ne sont pas propres aux algorithmes comparés et donc utilisées dans les deux cas. Puis les images et les statistiques de nombre de rayons d'ombrage par pixel sont présentés pour une scène de test.

6.6.1 Algorithme de Monte Carlo adaptatif

L'algorithme de Monte Carlo adaptatif utilisé a été présenté par Kok et Jansen en 1992 [KJ92]. Son principe est le suivant : les échantillons sont distribués sur une source rectangulaire en utilisant un arbre binaire dont les divisions se font alternativement suivant les directions des côtés du rectangle. Les noeuds de l'arbre obtenu sont des sous-rectangles de la source. Pour chaque noeud, la position de l'échantillon vers lequel le rayon est lancé est tirée au hasard. Quand un noeud est subdivisé, son échantillon est transmis à celui des sous-noeuds qui le contient. Sa visibilité est ainsi réutilisée sans avoir à lancer de rayon supplémentaire.

Pour calculer l'ombrage d'un point, l'algorithme commence par lancer un nombre de rayons initiaux qui dépend de l'énergie qui peut être reçue potentiellement puis estime la variance des échantillons. Si celle-ci dépasse un certain seuil, de nouveaux échantillons sont ajoutés. La procédure de choix de la position d'insertion d'un nouvel échantillon descend de la racine jusqu'à une feuille de l'arbre en utilisant des heuristiques pour choisir les sous-arbres. Des contraintes d'équilibre de l'arbre sont également imposées. Tous ces choix ont pour but de concentrer les échantillons le long du contour visible de la source.

6.6.2 Méthode de calcul des images

Le calcul des images procède ligne par ligne. Pour une ligne donnée, tous les rayons directs sont d'abord lancés ce qui permet de segmenter la ligne selon l'objet qui est visible. Pour chaque segment de pixels contigus, on construit la boîte englobante des points d'intersection des rayons dans la scène. Cette boîte contient donc tous les points où l'ombrage va être calculé et est utilisée pour construire une liste d'obstacles candidats grâce à la méthode de *shaft culling* de Haines et Wallace [HW91]. Nous utilisons aussi une hiérarchie de boîtes englobantes qui est construite en utilisant la méthode de Goldsmith et Salmon [GS87]. Dans [HW91], le chiffre d'une dizaine de rayons est donné comme seuil de rentabilité pour la construction de la liste de candidats, valeur qui dans notre cas est toujours

atteinte pour un segment de quelques pixels. De plus, si la liste est vide, il suffit d'intégrer le long du bord de la source et aucun rayon n'est nécessaire.

Dans le calcul du contour approché, la subdivision récursive des cellules indique la complication du contour de la partie visible de la source. Le raffinement récursif permet de traiter ce problème mais n'est pas forcément très efficace si le nombre de rayons primaires est vraiment insuffisant. Pour le calcul de l'éclairage pixel après pixel, la profondeur de raffinement est un bon indicateur de la complexité du contour intégré. Dans notre implémentation, nous l'utilisons pour faire varier adaptativement le nombre de rayons primaires. La règle employée est la suivante : si la subdivision récursive est descendue de plus d'un niveau dans l'arbre, on augmente le nombre de rayons primaires en commençant un niveau plus profond au pixel suivant, si la subdivision récursive est descendue d'un seul niveau, on ne fait rien et si il n'y a pas eu de raffinement, on diminue le nombre des rayons primaires en remontant d'un niveau dans l'arbre.

Une autre utilisation de la cohérence a été faite dans le cas de l'intégration de contour approché. Au lieu de calculer l'éclairage de chaque pixel, on commence par ne lancer que les rayons de la grille régulière sans commencer la construction du contour approché. On obtient pour chaque pixel, un motif des rayons interceptés et des rayons atteignant la source. Entre deux pixels voisins, le changement d'état d'un rayon (intercepté/non intercepté) permet de segmenter encore plus finement la ligne de pixel. Du point de vue des discontinuités de l'éclairage, cette technique permet de trouver dans l'image, les événements de type EV pour lesquels le sommet appartient à la source. En effet, ces événements correspondent au changement de visibilité des sommets de la source vers lesquels des rayons primaires sont lancés. Les autres discontinuités de type EV ou celles de type EEE ne sont pas détectées. Pour chaque segment, on ne calcule alors que trois facteurs de forme, deux aux extrémités et un au centre. Les valeurs des autres pixels sont reconstruites en utilisant un schéma quadratique d'interpolation. Sur les objets courbes, c'est le vecteur de forme qui est interpolé puis multiplié par la vraie normale au centre du pixel. Le fait que l'on ne détecte pas toutes les discontinuités peut être gênant puisqu'on utilise ensuite l'interpolation. En pratique, si le nombre de rayons primaires est suffisant, les pénombres sont découpées en segments suffisamment petits pour éliminer les artefacts.

6.6.3 Comparaison des deux méthodes

La scène de test utilisée pour comparaison est schématisée à la figure 6.9. Elle comprend deux sphères et un rectangle à l'intérieur d'un cube, éclairés par une source rectangulaire. La petite sphère se trouve dans la pénombre de la grande. Des images de cette scène ont été calculées à une résolution de 800×800 avec un rayon par pixel. Les résultats pour différentes images sont résumés dans le tableau 6.1. Comme expliqué dans la section précédente, les optimisations utilisées permettent pour certains groupes de pixels de décider à l'avance qu'aucun

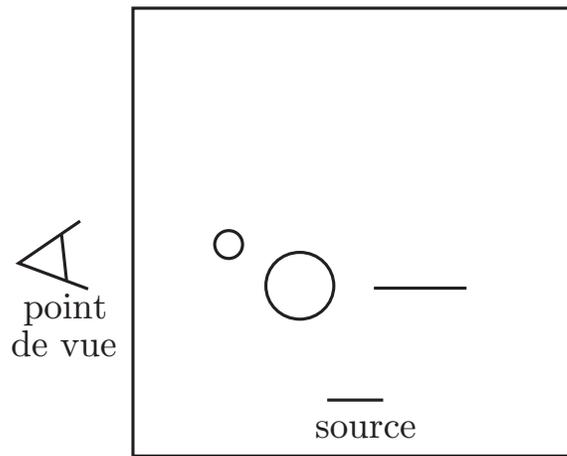
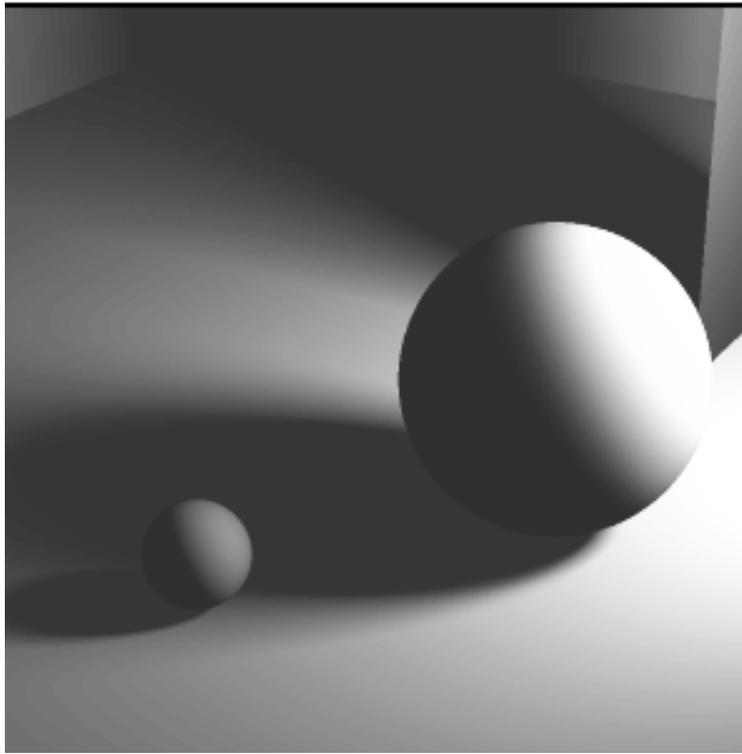


FIG. 6.9 - Schéma de la scène de test vue de dessus.

rayon d'ombrage n'est nécessaire car il n'y a pas d'obstacle candidat. Dans la deuxième colonne du tableau (nombre de rayons d'ombrage par rayon de vue), le calcul a été fait en se restreignant aux pixels pour lesquels des rayons d'ombrage sont effectivement lancés. La différence moyenne avec l'image de référence est calculée sur les valeurs des pixels qui sont comprises entre 0 et 255. Les temps de calcul sur une HP 710 sont indicatifs mais d'une manière générale, les algorithmes de choix des échantillons (pour le contour approché ou la méthode de Monte Carlo) ajoute un surcoût peu important au coût du lancer des rayons d'ombrage. Par exemple, même sur la scène de test très simple, la part de l'ensemble des fonctions liées à l'intégration de contour ne représente que 5 % du temps total de calcul. Pour des scènes plus complexes où le coût pour lancer un rayon est plus important, cette part devient négligeable.

Les différentes images calculées sont :

- image de référence calculée en utilisant 1024 rayons d'ombrage par pixel ;
- image obtenue par intégration de contour approché avec 9 échantillons de départ sur la source et 5 étapes de dichotomie en calculant le facteur de forme complet pour chaque pixel ;
- image obtenue par intégration de contour approché avec 9 échantillons de départ sur la source et 5 étapes de dichotomie en utilisant l'interpolation quadratique décrite au §6.6.2 ;
- image obtenue par la méthode de Monte Carlo en ajustant le seuil de variance pour obtenir environ le même nombre de rayons d'ombrage que dans l'image précédente ;
- image obtenue par la méthode de Monte Carlo en ajustant le seuil de variance pour réduire le bruit à un niveau acceptable.

FIG. 6.10 - *Image de référence.*

La figure 6.10 montre l'image de référence calculée à la résolution de 800×800 . La figure 6.11 montre un même détail pour les différentes images calculées.

6.7 Conclusion

La technique d'intégration de contour approché calcule un éclairage précis et non bruité aussi bien dans les zones de pleine illumination que dans les pénombres. La différence moyenne avec la solution de référence est inférieure à un quart de niveau d'intensité. La méthode se compare favorablement avec une méthode de Monte Carlo adaptative. Cependant, à mesure que le contour devient plus complexe, il est clair que la méthode de Monte Carlo prend l'avantage sur l'intégration de contour. La profondeur de raffinement qui est utilisée pour détecter la complication du contour pourrait servir de base à une méthode hybride capable de changer d'algorithme en cours de calcul d'une image.

Une remarque intéressante peut être faite en observant le nombre moyen de rayons lancés par pixel pour calculer l'image 3. Ce nombre 11.4 est très peu supérieur au nombre des rayons primaires (9 dans ce cas) utilisés à chaque pixel pour segmenter la ligne avant interpolation. Le surcoût lié au calcul de l'éclairage une fois que ses discontinuités sont localisées est faible. Nous avons déjà vu que

Image	Rayons d'ombrage	Rayons d'ombrage par rayon direct	Av. Dif.	Temps
1	603,456,646	1024.0	-	-
2	13,018,921	26.2	0.22	10 min
3	5,973,391	11.4	0.23	4 min 30
4	6,084,145	11.6	1.31	7 min 30
5	18,647,345	35.6	0.64	15 min

TAB. 6.1 - *Statistiques pour les images de test.*

les rayons testant la visibilité avec les sommets de la source permettent de trouver certaines des discontinuités de type EV. Dans le chapitre suivant nous montrons comment trouver les autres et réduire encore le temps nécessaire au calcul de l'éclairage par une source étendue.

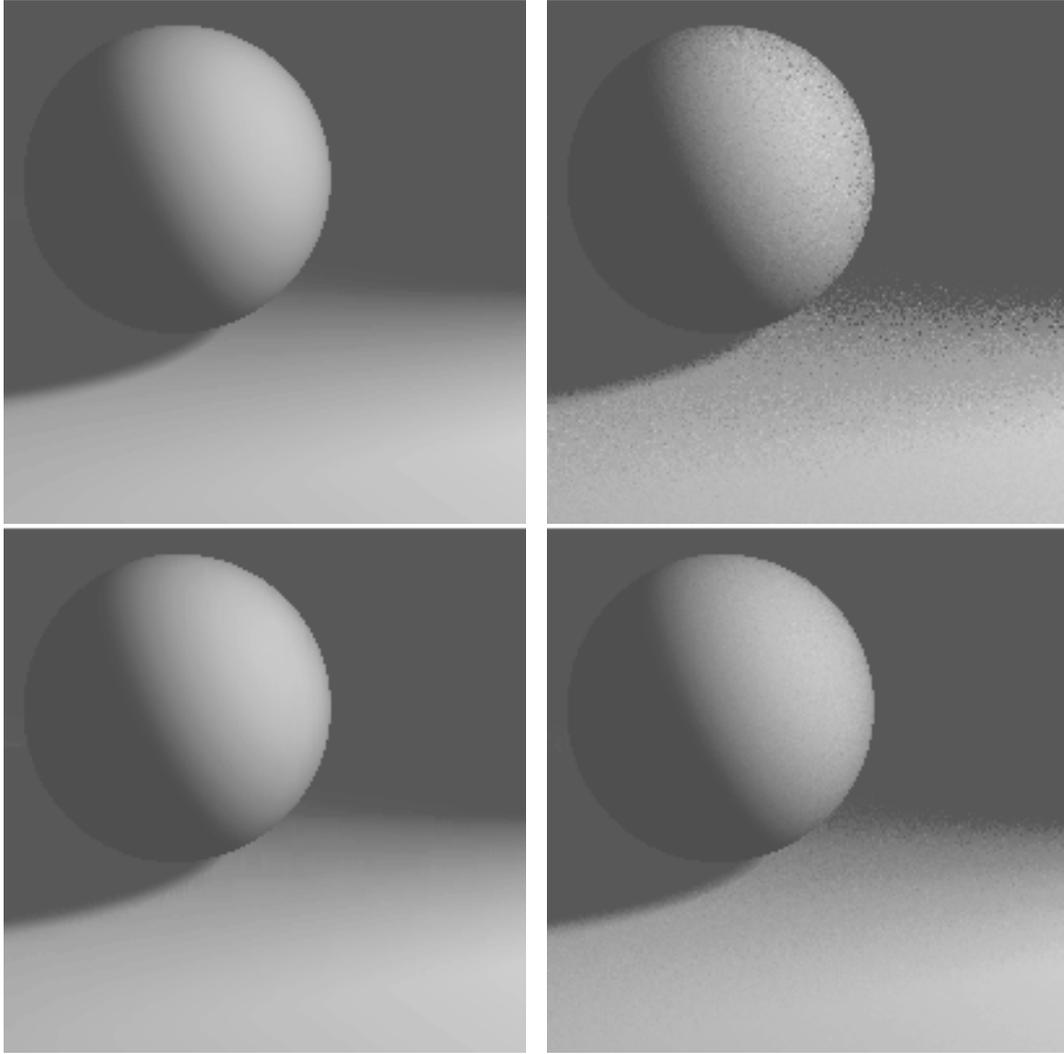


FIG. 6.11 - *Détail des quatre images. Images 1, 3, 4 et 5, de haut en bas et de gauche à droite.*

Chapitre 7

Maillage discret des discontinuités pour des surfaces courbes

Dans ce chapitre, nous étendons l'utilisation des discontinuités pour la reconstruction de l'illumination aux surfaces courbes. Les différents types de discontinuités possibles sont d'abord caractérisés. Puis un algorithme de découpage de l'image discrète, selon certaines de ces discontinuités est proposé, qui permet des calculs accélérés en réduisant l'ensemble de points où est évalué l'éclairage.

7.1 Discontinuités de l'illumination sur des surfaces courbes

7.1.1 Introduction

Quand une surface d'intensité lumineuse uniforme éclaire un environnement, les discontinuités de la fonction de radiosit  sur la surface des objets proviennent de changements qualitatifs de la portion visible de la source. Ces changements, en fait les singularit s de la projection de ces surfaces, ont d j   t   tudi s pour des applications autres que la synth se d'image. Dans le domaine de la vision par ordinateur, o  on les appelle  v nements visuels, ils sont utilis s pour faire une partition de l'ensemble des points de vue en zones dans lesquelles la silhouette d'un objet est qualitativement stable. Ils permettent ainsi de caract riser les diff rents aspects d'un objet. En reprenant des travaux de Gigus et Malik sur l'aspect des poly dres [GM90], il a  t  possible de faire l'inventaire des discontinuit s dans un environnement compos  de polygones [Hec91]. Elles appartiennent aux types EV et EEE dont nous avons parl  au chapitre 2. Nous proposons d' tendre ce r sultat   des objets lisses par morceaux en reprenant une classification de leurs  v nements visuels due   Rieger [Rie87].

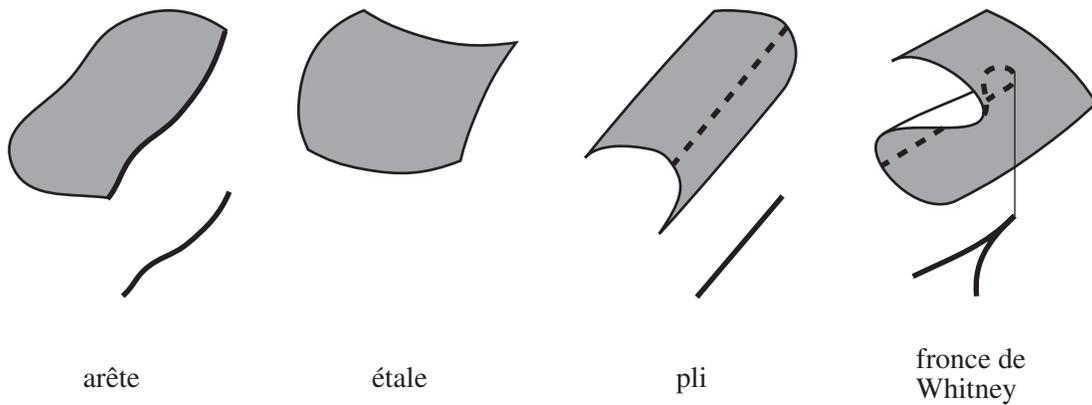


FIG. 7.1 - Les singularités de codimension 0 pour des objets lisses par morceaux.

Plus précisément, nous considérons donc des objets formés de morceaux de surfaces (zéros de fonctions C^∞ , $f_i(x, y, z)$) et de leurs courbes d'intersections supposées lisses elles aussi). Ces conditions ne sont pas très restrictives et sont satisfaites par beaucoup d'objets utilisés habituellement en synthèse d'image (par exemple, construits par *constructive solid geometry* ou assemblés à partir de carreaux paramétriques recollés continûment). Notre but, dans l'étude des discontinuités, est de permettre un meilleur « maillage » des surfaces, nous nous concentrons donc sur celles qui se produisent le long de courbes tracées sur les objets. Ces courbes correspondent à des surfaces dans l'espace à trois dimensions des points de vue pour la projection conique, c'est-à-dire à des singularités de codimension 1. Il existe également des singularités de codimensions 0 et supérieures à 1. Les secondes ne nous concernent pas car elles se produisent le long de courbes dans l'espace (codimension 2) et donc en des points isolés des surfaces, ou en des points isolés de l'espace (codimension 3) et donc pas de manière générique sur les surfaces des objets. Par contre, nous retrouverons les singularités de codimension 0 dans la suite. Ce sont les caractères de la projection d'un objet qui restent stables quand on déplace le point de vue dans un voisinage de l'espace. Elles sont, pour les objets qui nous intéressent de quatre types distincts : l'arête, l'étale, le pli et la fronce de Whitney (voir figure 7.1). Les arêtes correspondent aux courbes d'intersection entre les différents morceaux de surface composant l'objet. Les trois autres types se définissent à partir du contour d'occlusion qui est, pour un point de vue donné, la courbe des points où la ligne de vue est tangente à la surface de l'objet. L'étale est l'image de l'objet hors du contour d'occlusion. Le pli est l'image du contour d'occlusion lui-même. La fronce ajoute la condition supplémentaire de tangence entre la ligne de vue et le contour et donne un point de rebroussement de la silhouette.

7.1.2 Classification

Pour les objets lisses par morceaux, Rieger dénombre treize types différents de singularités de codimension 1. Quatre de ces singularités sont dites locales, c'est-à-dire qu'elles sont images du voisinage d'un point unique sur la surface. Les neuf autres sont multilocales, et correspondent à la superposition de plusieurs singularités distinctes de codimension 0 en un même point de projection.

Singularités locales

Elles sont au nombre de quatre: la fronce (point de rebroussement d'une arête), la queue d'aronde, la lèvres et le bec à bec. Du point de vue des discontinuités d'illumination, ces singularités qui mettent en jeu un seul objet ne peuvent se produire qu'à condition que la source lumineuse soit cet objet. Si on se limite à des sources convexes (ou polygonales comme on le fera par la suite), ces quatre possibilités sont donc exclues.

Singularités multilocales

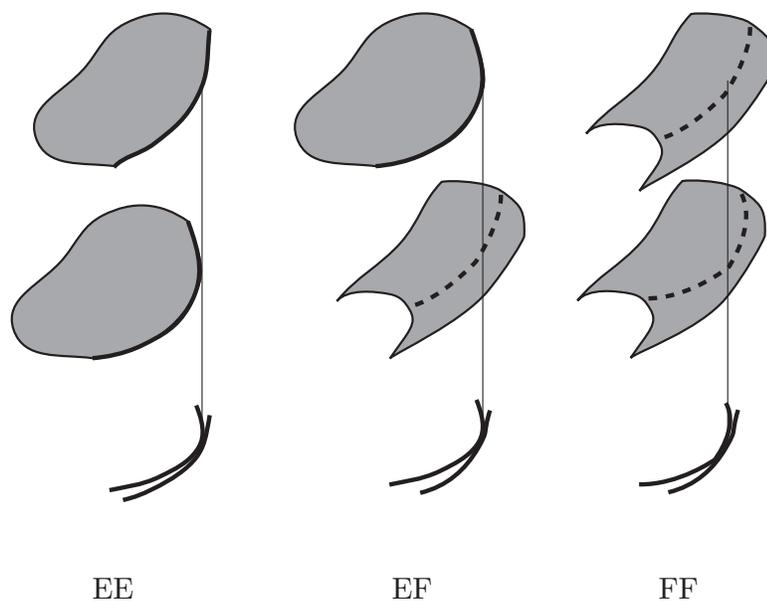


FIG. 7.2 - Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types *EE*, *EF* et *FF*).

Le nombre élevé de ces singularités provient des combinaisons possibles entre les types de silhouette (arête, pli ou fronce) et de contact le long de la ligne de vue (intersection ou tangence). Les neuf singularités multilocales sont présentées aux figures 7.2, 7.3 et 7.4. On utilise pour les nommer la lettre E pour désigner

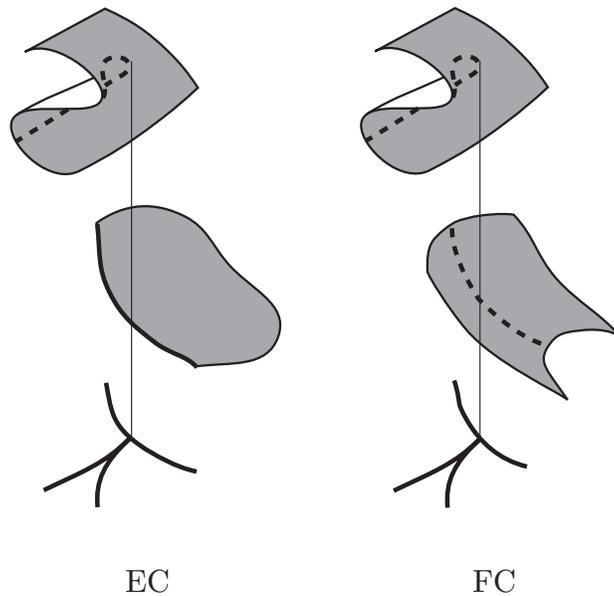


FIG. 7.3 - Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types EC et FC).

une arête (*edge* en anglais), la lettre F pour désigner un pli (*fold*) et la lettre C pour désigner une fronce de Whitney (*cusp*). On retrouve ainsi le type EEE de Gigus et Malik.

Comme dans le cas des singularités locales, les événements de type FC et EC ne peuvent pas produire de discontinuité d'illumination si la fronce de Whitney n'est pas produite par la source de lumière elle-même. En effet, cette fronce n'est visible que depuis un seul côté de l'objet qui la porte. Restent donc deux types de discontinuités suivant le nombre d'objets impliqués, les points triples : EEE, FFF, EFF et EEF d'une part et les points de tangence : EE, FF et EF d'autre part. En pratique, le calcul des surfaces de point triple est difficile, hormis le cas de trois arêtes droites qui donne une surface quadrique. Dans la suite, on n'utilisera pas ces événements dans le découpage. La propriété qui permet de les ignorer est qu'elles ne peuvent se produire que strictement à l'intérieur de la pénombre et donc qu'elles ne forment jamais la limite entre la zone de pleine illumination et la pénombre. En effet, soit a , b et c les trois voisinages sur les objets ainsi ordonnés le long de leur tangente commune au point triple. Pour qu'il y ait discontinuité de l'illumination, il faut que la source soit derrière les deux autres objets. On se trouve donc forcément dans la situation de l'une des deux configurations de la figure 7.5 à l'échange de b et c près. La source a est donc partiellement cachée de part et d'autre de la singularité.

Dans la suite, nous allons choisir une source polygonale. La situation est un peu différente puisque les arêtes ne sont pas lisses mais possèdent des sommets. Il faut donc rajouter les cas de coïncidence d'un sommet et d'une singularité

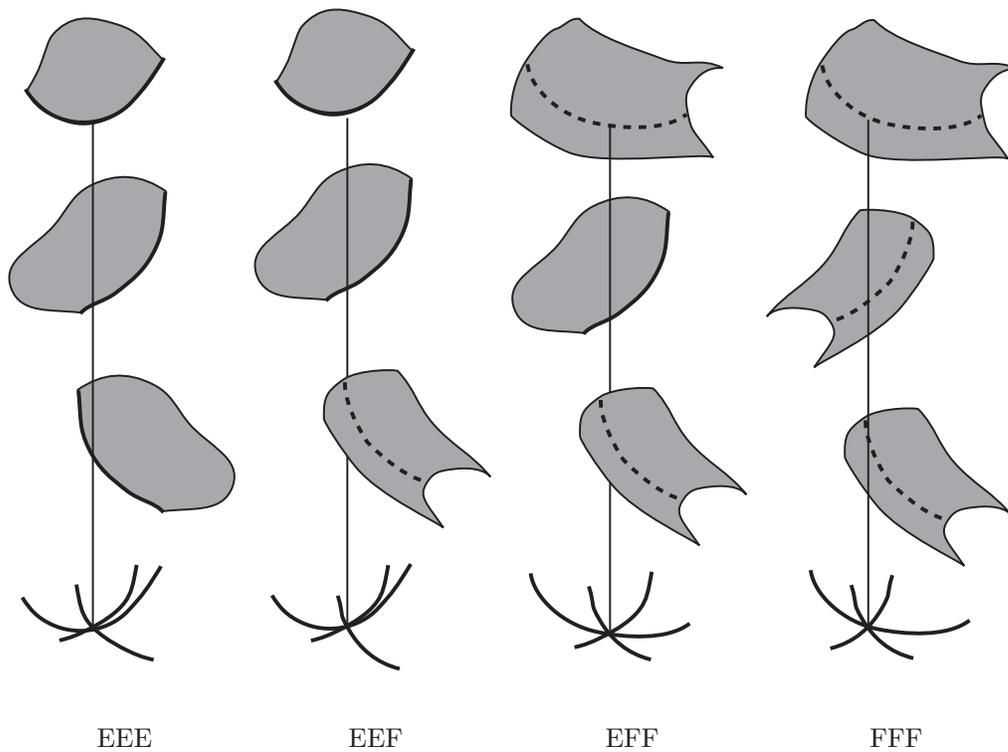


FIG. 7.4 - Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types *EEE*, *EEF*, *EFF* et *FFF*).

de codimension 0 (arête ou pli). Dans le nom de la discontinuité, on réserve la première lettre à la source et la seconde à l'autre objet impliqué. On distingue donc les cinq cas : EF (arête de la source et pli d'un objet), EE (arête de la source et arête d'un objet), EV (arête de la source et sommet d'un objet), VE (sommet de la source et arête d'un objet) et VF (sommet de la source et pli d'un objet).

D'un point de vue pratique, deux groupes se séparent. EF, EE et EV (désignés par E^*) donnent comme discontinuités des secteurs angulaires plans. En effet, la traversée de l'arête ne peut se faire que par un point isolé ou par un segment parallèle à l'arête. VE et VF (désignés par V^*) donnent des cônes issus du sommet et s'appuyant sur les objets. Ils seront tous les deux détectés par la même méthode.

7.1.3 Principe de l'algorithme

Dans le chapitre précédent, nous avons vu que la détection de certaines discontinuités permet de réduire le nombre de points où l'illumination est évaluée et donc le nombre de rayons d'ombrage nécessaires. Pour y arriver, on lance quelques rayons (les rayons primaires de l'intégration de contour approché) vers des points de la source et on groupe les pixels successifs pour lesquels la visibilité de ces

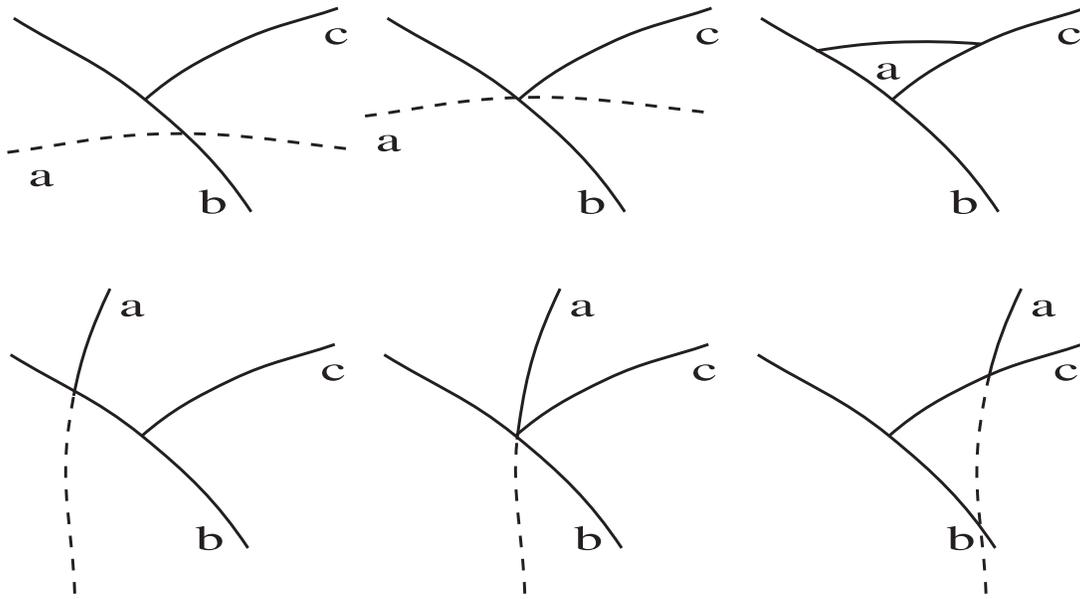


FIG. 7.5 - Passages d'une singularité multilocale de type point triple.

points est stable. Dans ce découpage, le changement de visibilité des sommets correspond au passage des discontinuités de type V^* (sommet de la source avec bord de l'obstacle). Les autres changements, associés à des points placés strictement à l'intérieur des arêtes ou de la source elle-même, ne sont pas associés à des discontinuités et on peut donc penser pouvoir les éliminer. La méthode décrite à la fin du chapitre 6 ne permet pas cette élimination car l'oubli des autres types de discontinuités (points triples et surtout E^*) interdit des segments d'interpolation trop longs. Ce sont les points ajoutés sur la source en plus des sommets qui entraînent automatiquement le découpage plus fin nécessaire.

Nous venons de voir que même en présence d'objets courbes et pourvu que la source soit polygonale, les surfaces de discontinuité dues à une arête de la source sont simples. Ce sont des secteurs angulaires plans s'appuyant sur les objets. Les discontinuités de type point triple sont beaucoup moins faciles à calculer mais comme elles sont situées dans l'intérieur des pénombres, on peut (comme dans [LTG92] pour des scènes polyédriques) les ignorer à condition de tester ensuite la validité de chaque interpolation faite dans la pénombre et de la raffiner si nécessaire.

Nous proposons de construire un algorithme de calcul d'image tenant compte des discontinuités de types EF, EE, EV, VE et VF pour segmenter l'ensemble des pixels de l'image et utilisant l'interpolation dans les zones obtenues pour ne calculer l'illumination qu'en un nombre réduit de pixels. Les principaux points de l'algorithme sont :

- calcul des discontinuités E^* et tri des plans supports obtenus dans une structure d'arbre binaire;

- détection des discontinuités entre deux pixels adjacents;
- segmentation des colonnes de pixels et regroupement de ces segments en bandes horizontales;
- interpolation de l'illumination à l'intérieur des bandes.

Ces différents points sont abordés en détail dans la suite. On montre également comment utiliser la structure de données contenant les discontinuités des arêtes de la source pour créer des listes de candidats pour le lancer de rayons en remplacement de la technique de *shaft culling* [HW91].

7.2 Algorithme de calcul d'image

7.2.1 Calcul et tri des discontinuités E^*

Les discontinuités E^* sont des secteurs angulaires. En pratique, la structure associée à une telle discontinuité comprend l'équation du plan π support du secteur et celles de deux ou trois demi-espaces définissant le secteur angulaire. Cette représentation permet, une fois connu le point de traversée du plan π de savoir rapidement si il appartient ou non au secteur. Pour une discontinuité EF , on recherche un point de l'objet où le plan tangent contient les deux extrémités A et B de l'arête. Dans le cas d'une surface implicite définie par l'équation $f(x, y, z) = 0$, on obtient le système de trois équations suivant :

$$\begin{cases} f(x, y, z) = 0, \\ (x - x_A)f_x(x, y, z) + (y - y_A)f_y(x, y, z) + (z - z_A)f_z(x, y, z) = 0, \\ (x - x_B)f_x(x, y, z) + (y - y_B)f_y(x, y, z) + (z - z_B)f_z(x, y, z) = 0. \end{cases}$$

Dans le cas d'une surface paramétrique $F(u, v)$, on obtient le système de deux équations :

$$\begin{cases} \left(\frac{\partial \vec{F}}{\partial u} \times \frac{\partial \vec{F}}{\partial v} \right) \cdot \overrightarrow{AF} = 0 \\ \left(\frac{\partial \vec{F}}{\partial u} \times \frac{\partial \vec{F}}{\partial v} \right) \cdot \overrightarrow{BF} = 0. \end{cases}$$

La résolution pratique de ces systèmes réclame encore du travail. Une résolution en deux étapes est envisageable. L'algorithme consisterait à construire le contour d'occlusion pour un des sommets de l'arête (comme par exemple dans [EC90]) puis à parcourir ce contour en testant l'orientation de la normale de l'objet par rapport à la direction vers l'autre sommet de l'arête. Cette deuxième étape serait aussi utilisée dans la recherche des discontinuités EE .

Enfin, aucun calcul n'est nécessaire pour les discontinuités EV qui sont définies par les extrémités de l'arête et un sommet d'un obstacle. Notre implémentation est limitée à des objets très simples (polygones et sphères) mais permet de montrer la

validité de la suite de l'algorithme. Pour des objets plus complexes, la complexité du calcul des événements EF/EE augmentent. C'est aussi le cas du coût de calcul d'intersection d'un rayon ce qui laisse la méthode rentable.

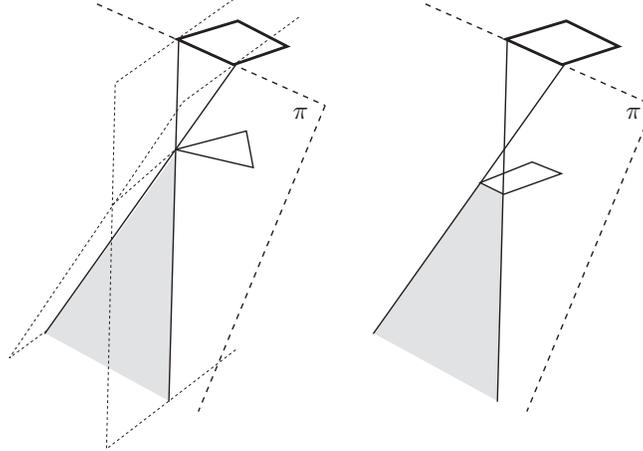


FIG. 7.6 - Secteurs angulaires de discontinuité VE pour un sommet (à gauche) et une arête parallèle à celle de la source (à droite). Le coin du secteur de droite est tronqué par le plan de l'obstacle.

Cas des polyèdres

Si l'objet est constitué de facettes polygonales, les discontinuités EE ne peuvent se produire que si l'arête de l'obstacle est parallèle à celle de la source. Le secteur obtenu doit alors être tronqué comme indiqué sur la figure 7.6 à droite.

Cas des sphères

Nous utilisons une procédure spéciale pour la sphère avec laquelle il n'y a que des discontinuités EF. Soit l'arête AB et une sphère \mathcal{S} de centre C et de rayon r . Les points M tels que le plan ABM est tangent à \mathcal{S} sont solutions du système :

$$\begin{cases} \overrightarrow{CM} \cdot \overrightarrow{AM} = 0 \\ \overrightarrow{CM} \cdot \overrightarrow{BM} = 0 \\ \overrightarrow{CM} \cdot \overrightarrow{CM} = r^2. \end{cases}$$

En combinant chacune des deux premières équations avec la troisième, on obtient :

$$\begin{cases} \overrightarrow{CM} \cdot \overrightarrow{CA} = r^2 \\ \overrightarrow{CM} \cdot \overrightarrow{CB} = r^2 \\ \overrightarrow{CM} \cdot \overrightarrow{CM} = r^2, \end{cases}$$

et en éliminant deux des coordonnées de M grâce aux deux premières relations (linéaires en M), on se ramène à résoudre une équation du second degré. On trouve zéro, une ou deux solutions suivant que la droite support de AB intersecte, est tangente à ou n'intersecte pas \mathcal{S} .

Tri des discontinuités

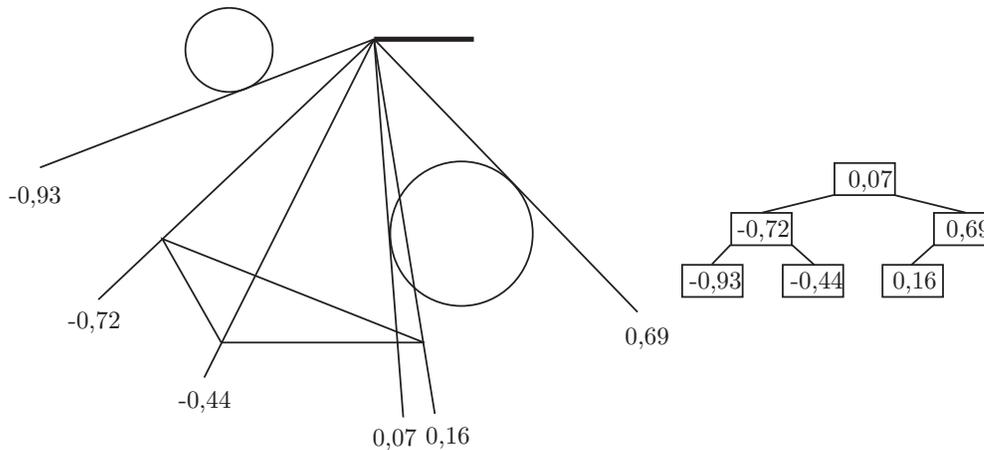


FIG. 7.7 - Tri des plans supports des discontinuités E^* autour de l'arête (en projection dans un plan normale à l'arête). A droite, l'arbre binaire équilibré associé dont les noeuds contiennent en plus l'équation du plan support et des demi-espaces définissant le secteur angulaire.

Pour une arête de la source, on obtient avec chaque objet une série de plans contenant chacun un secteur angulaire de discontinuité. Dans la suite, il faudra déterminer si un segment joignant les points correspondant à deux pixels adjacents de l'image traverse ces secteurs angulaires. Pour mettre à profit la cohérence (c'est-à-dire la proximité) entre les segments traités successivement, on trie les plans en utilisant l'angle fait avec le plan de la source comme clef (voir figure 7.7). En pratique comme seules les discontinuités placées devant la source sont intéressantes (aucune lumière n'est émise vers l'arrière), on peut remplacer l'angle par son cosinus, soit la composante de la normale unitaire au plan dans la direction du plan de la source perpendiculaire à l'arête.

On utilise, pour trier les plans, des arbres binaires de recherche équilibrés de type arbres rouges et noirs. Cette structure a une taille linéaire dans le nombre de plans introduits et permet de réaliser en temps logarithmique les opérations d'insertion, de localisation et de recherche de prédécesseur ou de successeur dans l'ordre qui seront utilisées dans la suite (voir figure 7.7).

7.2.2 Détection des discontinuités

On dispose au départ d'un segment de pixels p_0, \dots, p_n , adjacents couvrant un même objet intersecté aux points P_0, \dots, P_n et éclairés par la source lumineuse. L'algorithme de segmentation qui permet d'arriver à ce point sera vu dans la section suivante.

Discontinuités V^*

Les surfaces de discontinuité de type V^* sont des portions de cônes s'appuyant sur les contours d'occlusion ou les arêtes des obstacles et issus des sommets de la source. Plutôt que de tester les intersections avec ces cônes qui peuvent prendre des formes complexes, il est possible d'utiliser le lancer de rayons et de bénéficier ainsi des méthodes d'accélération qui existent. Le passage au travers d'une surface V^* marque le changement de visibilité d'un sommet de la source. Il suffit donc, à chaque pixel p_i , de lancer un rayon depuis P_i vers chacun des sommets pour repérer ces surfaces.

Discontinuités E^*

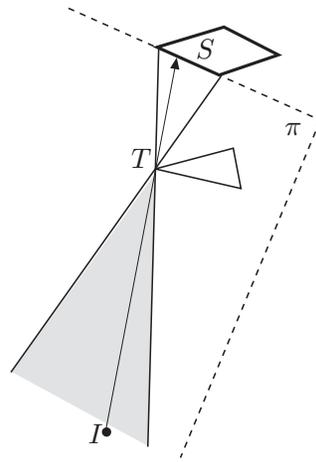


FIG. 7.8 - Rayon \vec{IS} utilisé pour vérifier l'existence réelle de la discontinuité EV .

Les surfaces de discontinuité de type E^* sont de simples secteurs angulaires plans. Nous les avons calculés explicitement et triés autour de l'arête. Pour détecter les intersections avec ces secteurs efficacement, on calcule la clef k_0 du point P_0 (par l'intermédiaire de la normale au plan contenant l'arête et P_0). Puis on localise, dans l'arbre binaire de recherche de l'arête, les plans de discontinuité dont les clefs K_j et K_{j+1} sont respectivement immédiatement inférieure et supérieure à k_0 . On considère ensuite le pixel p_1 et on compare sa clef k_1 à K_j et K_{j+1} . Trois cas se présentent : $K_j < k_1 < K_{j+1}$, $k_1 \leq K_j$ et $K_{j+1} \leq k_1$. Dans le

premier cas, on ne peut pas avoir traversé un plan de discontinuité ni *a fortiori* un secteur angulaire. Les deux autres cas sont symétriques, supposons par exemple que $K_{j+1} \leq k_1$. On a donc traversé le plan support π_{j+1} de la discontinuité de clef K_{j+1} . On calcule alors le point d'intersection I du segment P_0P_1 avec ce plan et on teste si ce point se trouve à l'intérieur ou bien à l'extérieur du secteur angulaire de discontinuité. Dans le premier cas, il reste une dernière vérification à effectuer pour savoir si on a bien affaire à une discontinuité E^* : il faut que le sommet T du secteur angulaire et le point S correspondant sur l'arête soient tous deux visibles depuis I . Il faut donc lancer un rayon de I vers S jusqu'à la source pour vérifier qu'il n'y a pas d'obstruction (voir figure 7.8). Le rayon passant très près de l'objet responsable de la discontinuité au niveau de T , il faut s'assurer qu'on n'intersecte pas cet objet au voisinage de T ce qui constituerait un « faux obstacle ».

Dans le deuxième cas, on cherche le plan π_{j+2} suivant π_{j+1} dans l'ordre en temps logarithmique, grâce à la structure d'arbre binaire équilibré. Puis on recommence l'opération précédente jusqu'à trouver une discontinuité ou un nouvel encadrement valide de la clef k_1 entre les clefs de deux plans successifs de l'arbre. On passe ensuite au pixel suivant.

7.2.3 Segmentation des colonnes de pixels

Pour chaque colonne de pixels, on commence par lancer tous les rayons directs et on stocke dans un tableau chaque point d'intersection avec sa normale et un pointeur sur l'objet intersecté. Parallèlement à ce tableau, on maintient une liste doublement chaînée de segments qui décrit le découpage de la colonne. L'information associée à chaque segment se compose :

- d'une étiquette décrivant le type du segment;
- des valeurs de visibilité des sommets de la source;
- de pointeurs permettant le chaînage vertical (colonne) et horizontal (bande)

Au départ de l'algorithme, un segment unique recouvre toute la colonne. On découpe ensuite ce segment initial en effectuant dans l'ordre les cinq tests suivants :

1. appartenance à un même objet (les rayons n'ayant pas rencontré d'objet sont groupés en segments de type FOND, les autres segments sont laissés sans type défini);
2. position par rapport au plan de la source (les pixels situés à l'arrière de la source sont groupés en segments de type ARRIÈRE, les autres segments sont laissés sans type défini);

3. position des sommets de la source par rapport à l'horizon du point (le découpage se fait à chaque passage de l'horizon par un sommet. Les segments pour lesquels tous les sommets sont cachés sont marqués du type AUTO-OMBRAJE, les autres segments sont laissés sans type défini);
4. visibilité des sommets situés au dessus de l'horizon. Ce critère réalise le découpage suivant les discontinuités V^* . Tous les segments sont laissés sans type défini;
5. traversée des discontinuités E^* . Là encore, on conserve tous les segments.

A la fin de chaque étape, les segments d'un type défini sont éliminés de la liste et le critère suivant ne leur est donc pas appliqué. Dans tous les cas, il n'y a pas d'illumination à calculer (FOND) ou bien celle-ci est nulle (ARRIÈRE et AUTO-OMBRAJE). Après le troisième critère, on obtient des segments composés de points situés sur un même objet, devant la source et voyant les mêmes sommets de la source au dessus de leurs horizons. On peut donc appliquer les procédures 4 et 5 telles qu'elles ont été décrites au §7.2.2.

Le découpage suivant les discontinuités V^* (visibilité des sommets) se déroule de la manière suivante: on divise d'abord le segment entré en sous-groupes de pixels. Puis, on calcule la boîte englobante des points du sous-groupe ce qui permet de former une liste d'obstacles potentiels pour le lancer de rayons vers la source. Les sous-groupes de pixels dont la liste de candidats est vide sont coupés du reste et traités comme des segments à part entière bien qu'ils ne correspondent pas à de vraies discontinuités. Le calcul de l'éclairage de ces segments marqués du type ILLUM se fait par intégration de contour le long des arêtes de la source sans avoir à tester la visibilité. Ces segments sont conservés jusqu'à la fin de la segmentation. Pour chaque pixel des autres sous-groupes, on lance autant de rayons qu'il y a de sommets au dessus de l'horizon.

Après l'étape 5, on procède à l'extension des segments ILLUM. En général, la division en sous-groupes pour le calcul des listes de candidats ne correspond pas aux bords exacts des pénombres. On va donc trouver, au bord des pénombres, des segments sur lesquels tous les sommets de la source sont visibles mais qui ne sont pas pour autant reconnus de type ILLUM. Pour récupérer ces segments et les traiter plus simplement, on parcourt les segments de type ILLUM et on considère ceux de leurs voisins qui voient tous les sommets. Le type ILLUM ne peut pas être transmis directement à ces voisins car le découpage entre les deux segments peut avoir été fait avant l'étape 5 et, par un hasard malheureux, correspondre à une discontinuité de type E^* . On effectue donc un test de traversée de ces discontinuités au passage entre ces deux segments avant la transmission éventuelle du type ILLUM.

On élimine aussi les segments de pleine illumination qui n'ont pas de voisin de type ILLUM, par exemple parce qu'ils sont au bord d'un objet. Il faut les distinguer des segments qui voient tous les sommets de la source mais sont dans la pénombre

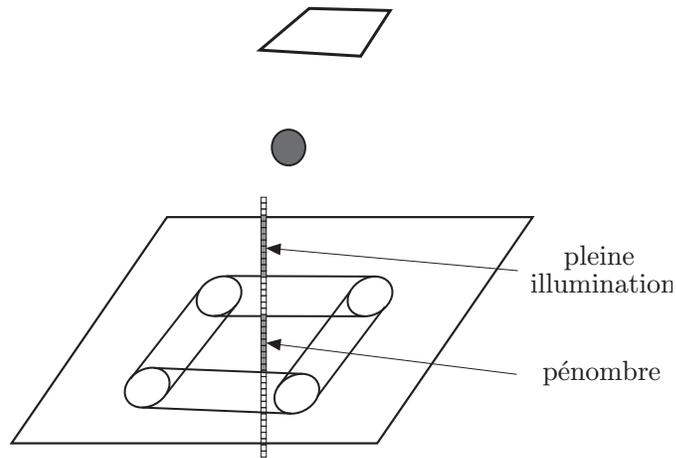


FIG. 7.9 - Deux types différents de segment d'où tous les sommets de la source sont visibles.

(voir figure 7.9). Pour cela, on calcule l'illumination au centre du segment par intégration de contour approché et si aucun obstacle n'est rencontré, on étend la propriété de pleine illumination à tout le segment qui est ensuite traité comme si il était de type `ILLUM` puis éliminé. Toutes ces extensions sont valides car les seules discontinuités restant alors non détectées sont de type point triple et ne peuvent donc pas se produire si aucun sommet de la source n'est caché.

En fin de segmentation, il ne reste que des segments dans l'ombre ou la pénombre. Nous décrivons d'abord la structure de données utilisée pour organiser les colonnes successives de pixels puis l'algorithme de regroupement des segments appartenant à des colonnes voisines en bandes.

7.2.4 Groupement de segments en bandes

Pour pouvoir interpoler l'illumination dans les deux dimensions du plan image et réduire ainsi le coût des calculs, il faut pouvoir conserver plusieurs colonnes successives de l'image. Ces colonnes sont stockées dans une liste doublement chaînée. Chaque élément comprend :

- des pointeurs vers les colonnes de gauche et de droite;
- un pointeur sur le tableau de pixels de la colonne;
- un pointeur vers la liste de segments restant à traiter pour la colonne.

Après avoir segmenter la colonne C_n , on la rattache à la liste de colonnes existantes $[C_m, C_{n-1}]$ et on essaye de prolonger les bandes déjà créées. Une bande de segments est une liste doublement chaînée de segments voisins dans la direction horizontale (un segment par colonne) et possédant certaines propriétés

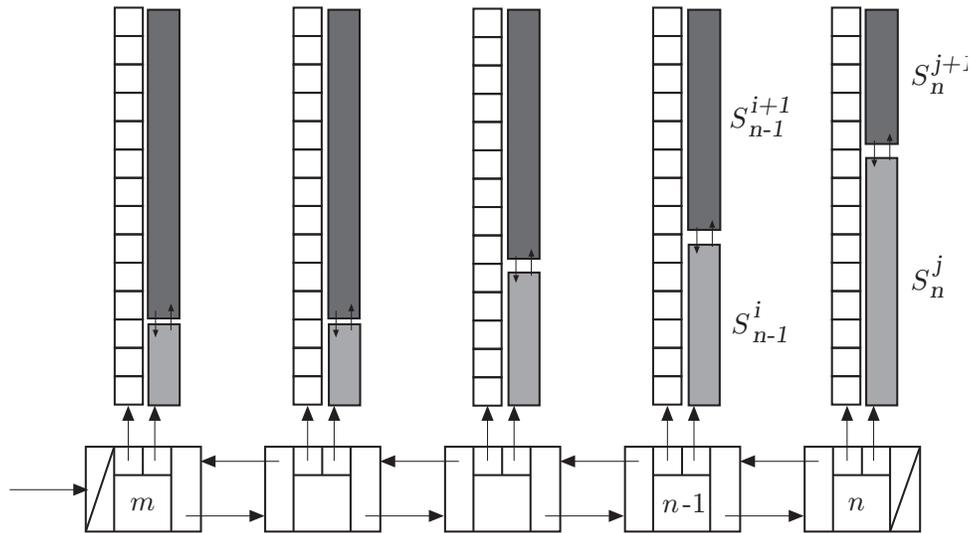


FIG. 7.10 - Organisation des colonnes de pixels et de segments pour permettre le rattachement en bandes de segments adjacents.

communes : même objet couvert, mêmes sommets de la source cachés et chevauchement entre deux segments voisins dans la direction verticale (voir figure 7.10).

Pour rattacher les segments de C_n aux bandes existant dans les colonnes de gauche, on parcourt C_{n-1} et C_n en parallèle. Pour chaque segment s_{n-1}^i de C_{n-1} , on parcourt les segments de C_n en vérifiant les propriétés énoncées plus haut jusqu'à trouver un bon segment ou à dépasser la hauteur de s_{n-1}^i . Deux cas sont alors possibles suivant que l'on trouve ou non un segment pour prolonger la liste :

1. s_n^j peut prolonger la bande de s_{n-1}^i . On met à jour les pointeurs de s_{n-1}^i et s_n^j et on passe à s_{n-1}^{i+1} en commençant les recherches dans C_n à s_n^{j+1} . Si des segments de C_n , s_n^{j-1} , s_n^{j-2} , etc. ont été passés avant de trouver s_n^j , ils démarrent de nouvelles bandes pour la suite;
2. aucun segment de C_n ne convient. La bande de s_{n-1}^i est terminée et transmise à l'algorithme d'interpolation décrit à la section suivante. On passe à s_{n-1}^{i+1} en reprenant la recherche dans C_n au premier segment libre.

Quand une bande est interrompue et traitée, tous ses segments sont effacés des listes des colonnes correspondantes. Après avoir rattaché tous les segments de C_n , on parcourt les premières colonnes de la liste C_m, C_{m+1}, \dots , jusqu'à en trouver une dont la liste de segments n'est pas vide. Les colonnes précédentes dont la liste est vide sont alors écrites dans le fichier image et libérées ainsi que le tableau de pixels correspondant. Cette méthode permet de ne conserver qu'une fenêtre de colonnes en mémoire à un instant donné. Si peu de mémoire est disponible, on peut décider de fixer à l'avance le nombre maximum de colonnes à garder

en mémoire. Pour cela, on garde à jour la longueur de chaque bande dans son segment le plus à droite, ce qui permet de terminer artificiellement la croissance d'une bande trop longue. On ne dépasse ainsi jamais le quota de colonnes en mémoire.

7.2.5 Interpolation de l'illumination

Pour calculer l'illumination d'une bande de pixels, on utilise deux interpolations successives suivant les directions horizontale et verticale. Le schéma employé est quadratique, c'est-à-dire que pour un segment, on calcule l'illumination au centre et aux extrémités et qu'on remplace celle des autres points par la valeur de la parabole passant par les trois valeurs calculées. Sur les surfaces courbes, c'est le vecteur de radiosité qui est interpolé ce qui permet d'utiliser ensuite la vraie normale au centre du pixel pour prendre le produit scalaire.

Interpolation horizontale

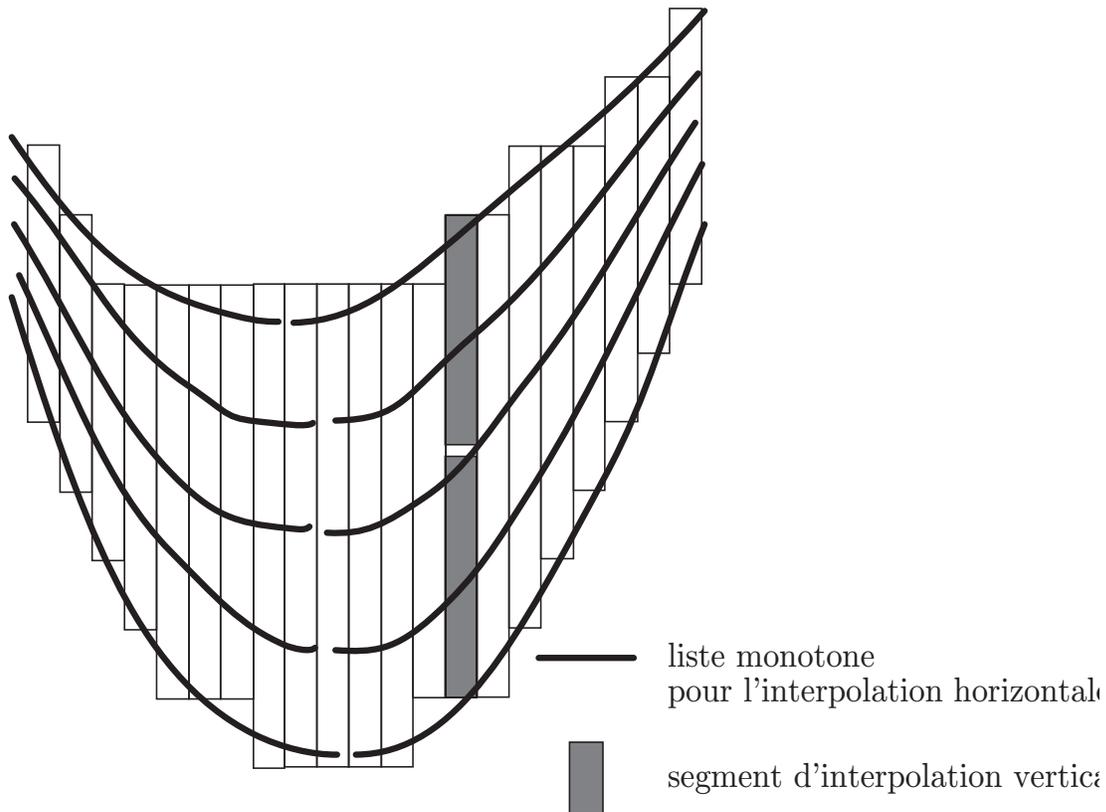


FIG. 7.11 - Lignes d'échantillons pour l'interpolation horizontale.

Pour pouvoir appliquer deux schémas quadratiques juxtaposés dans les colonnes, on calcule pour chaque bande, cinq listes de valeurs (voir figure 7.11).

```
si la longueur de la liste est inférieure a 5
alors
    calcul des points restants de la liste par intégration de contour.
sinon
    pour chacune des deux demi-listes
        calcul de la valeur au centre par intégration de contour.
        calcul de l'écart relatif entre cette valeur et celle obtenue
        par interpolation sur la liste entière.
        si l'écart relatif est trop grand
            alors
                appel récursif sur la demi-liste.
            sinon
                calcul des points restants par interpolation
```

FIG. 7.12 - *Algorithme d'interpolation horizontale.*

Pour une liste de points donnée, on commence par procéder à un découpage en sous-listes monotones (croissantes ou décroissantes). Pour chaque sous-liste, on calcule par intégration de contour approché les valeurs aux extrémités puis on applique l'algorithme récursif de la figure 7.12.

Interpolation verticale

Pour chaque segment, on dispose de cinq valeurs et on applique le schéma d'interpolation quadratique sur chaque demi-segment.

Placement des extrémités des segments et valeurs partagées

Le schéma d'interpolation décrit n'est valide que si la liste des points utilisés dans la première interpolation se trouve bien sur la discontinuité. Si on utilise directement les centres des pixels, l'interpolation verticale crée ensuite des discontinuités entre valeurs voisines. Le calcul plus précis du point de passage de la discontinuité entre deux pixels se fait au moment de la segmentation et est différent pour les deux types possibles de discontinuités :

- V*** l'intervalle dans lequel la visibilité du sommet change est réduit par dichotomie;
- E*** le point recherché est le point du plan support de la discontinuité que l'on a calculé pour vérifier qu'il appartenait bien au secteur angulaire.

Pour les discontinuités dans la dérivée de l'illumination (c'est-à-dire, dans l'image discrète, toutes sauf celles où tous les sommets changent de visibilité entre deux pixels voisins), on peut partager les valeurs entre deux bandes adjacentes dans le sens vertical et qui sont séparées par la discontinuité. Le partage se fait aussi avec les segments de type ILLUM qui avant d'être éliminés passent leur illumination aux segments de pénombre voisins. Pour cela, chaque structure de segment possède des champs contenant les points de passage précis des discontinuités à ses extrémités et les radiosités correspondantes. Avant d'appliquer l'interpolation, on vérifie donc que ces champs n'ont pas déjà été remplis lors du traitement d'une bande ou d'un segment voisin.

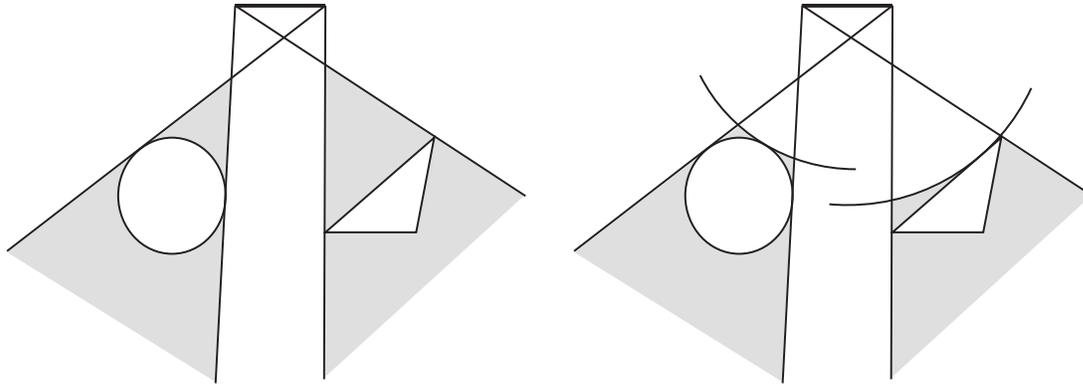


FIG. 7.13 - *A gauche, les volumes de pénombre obtenus en combinant les plans des discontinuités maximales pour chaque objet. Les zones hachurées placées devant les objets sont incorrectement classifiées. L'erreur est réduite à droite en ajoutant comme condition supplémentaire la distance au centre de la source.*

7.2.6 Candidats pour le lancer de rayons

Les plans de discontinuité E^* peuvent aussi être utilisés pour construire des listes réduites d'obstacles potentiels pour le lancer de rayons entre un groupe de points et la source. Cette méthode peut remplacer le *shaft culling* et son principe est le suivant. Moyennant un choix correct de l'orientation des discontinuités autour de l'arête, on note que le plan support de la plus grande discontinuité d'un objet donné le sépare de la source. Un point situé du même côté de ce plan que la source ne peut donc pas être dans l'ombre ou la pénombre de l'objet. En combinant les plans des discontinuités maximales de toutes les arêtes de la source, on obtient une pyramide contenant totalement le volume d'ombre et de pénombre de l'objet. On a supposé ici, pour plus de clarté, que la source est convexe ce qui permet d'utiliser toutes ses arêtes pour former la pyramide. Si ce n'est pas le cas, on peut toujours choisir un sous-ensemble acceptable des arêtes de la source. Pour les points situés plus loin de la source que l'objet, cette pyramide permet de

Image	Rayons d'ombrage	Rayons d'ombrage par rayon direct	Av. Dif.	Temps
1	603, 456, 646	1024.0	-	-
2	5, 973, 391	11.4	0.23	4 min 30
3	1, 952, 638	4.4	0.27	2 min 30

TAB. 7.1 - *Statistiques pour les images de la scène du chapitre 6. Référence (1), intégration de contour (2) et segmentation/interpolation (3).*

savoir assez précisément si l'objet est un obstacle potentiel puisqu'elle est directement tangente à l'objet plutôt qu'à une boîte englobante. Elle peut, par contre, retenir des candidats situés derrière les points par rapport à la source (voir figure 7.13). Pour limiter ce problème, on rajoute comme discriminant supplémentaire la distance au barycentre de la source.

Ce schéma peut être facilement relié à une hiérarchie de boîtes englobantes sur les objets de la scène. En effet le plan de la discontinuité maximum peut être codé par sa clef dans l'arbre binaire des discontinuités. Il suffit donc de rajouter à chaque noeud de la hiérarchie autant de nombres qu'il y a d'arêtes à la source plus un, la distance au centre de la source. On propage ces valeurs vers la racine de l'arbre en prenant le maximum des clefs de discontinuité pour chaque arête et le minimum des distances. La figure 7.14 illustre cette construction en dimension 2.

Pour construire la liste de candidats d'un groupe de points, on commence par calculer le volume englobant de ces points puis leurs clefs minimums pour chaque arête et leur distance maximum au centre de la source. On parcourt ensuite l'arbre de boîtes englobantes en élaguant les branches telles que l'une des clefs minimums est plus grande que la clef de même rang stockée au noeud ou telles que la distance maximum au barycentre est plus petite que celle stockée au noeud.

7.3 Résultats

L'algorithme décrit ci-dessus a été implémenté pour des scènes composées de polygones et de sphères et éclairées par des polygones. Les figures 7.15 et 7.16 illustrent sur un exemple les différents étapes de la segmentation décrites en 7.2.3.

Pour le tableau 7.1, on a repris la scène de test et, à la deuxième ligne, les résultats de l'algorithme du chapitre 6. La première ligne rappelle les données de l'image de référence. On observe une nouvelle réduction importante du temps de calcul grâce à l'algorithme de segmentation/interpolation (troisième ligne).

La figure 7.17 montre en gris les pixels de l'image de test où des rayons ont été lancés vers les sommets de la source et en noir ceux où le calcul d'interpola-

tion horizontale a été effectué. On note les cinq lignes différentes par bande de segments. Aucun rayon n'a été utilisé pour les pixels blancs. Nous rappelons que la troisième colonne du tableau 7.1 présentent des moyennes sur les pixels gris et noirs seuls.

Les résultats obtenus avec l'algorithme sont encourageants. Cependant, plusieurs points peuvent encore être développés :

- le calcul des discontinuités EF et EE pour d'autres types d'objets;
- l'utilisation d'un schéma d'interpolation adaptatif dans les deux dimensions;
- l'utilisation de la segmentation pour anti-aliasser les bords d'objets et les discontinuités importantes de l'illumination.

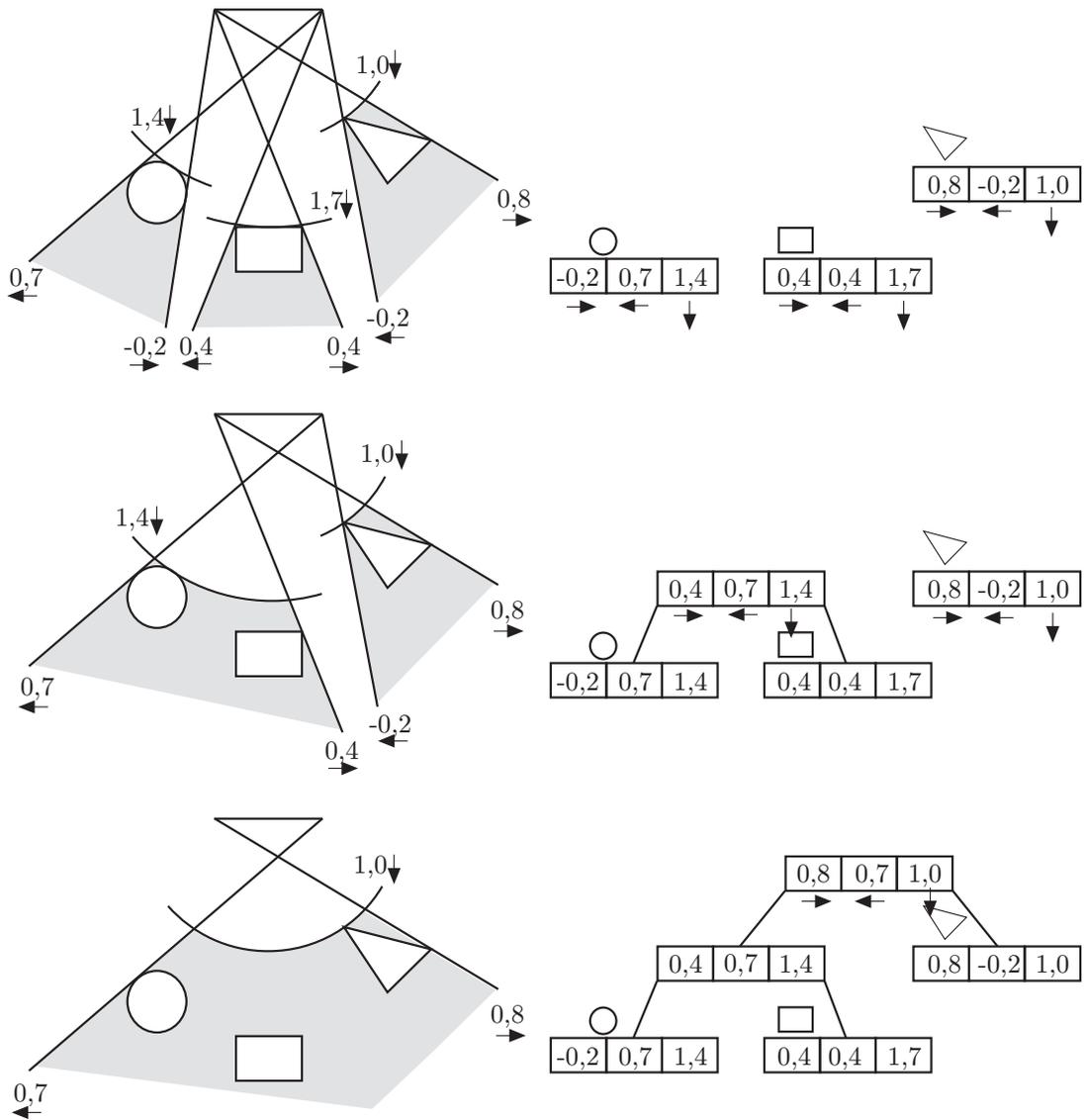


FIG. 7.14 - Propagation des discontinuités maximales dans la hiérarchie correspondant aux boîtes englobantes. Chaque noeud contient, en plus des informations habituelles non montrées, les clefs des discontinuité pour les arêtes de gauche et de droite et la distance minimale au centre de la source. Les secteurs de l'arête de droite sont orientés dans le sens des aiguilles d'une montre, ceux de l'arête de gauche en sens inverse.

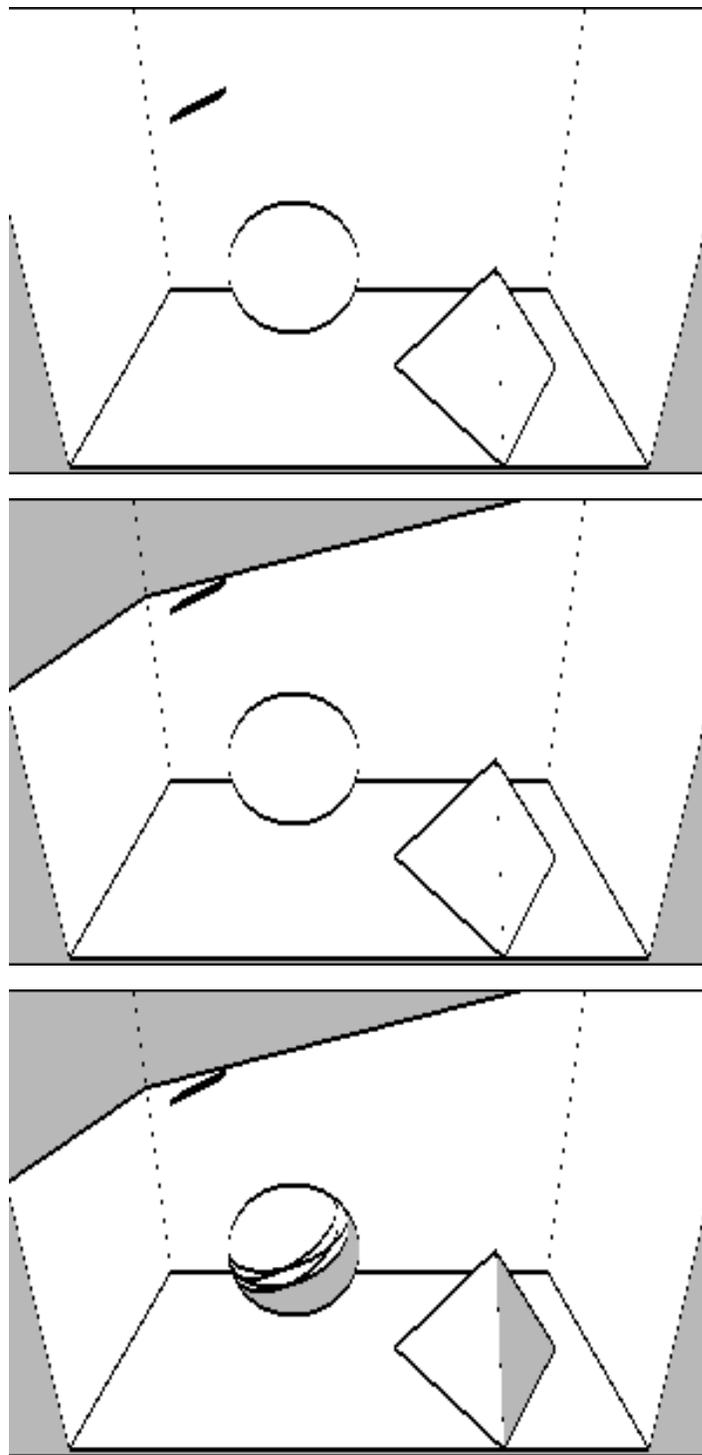


FIG. 7.15 - *Etapes de la segmentation. En haut, segmentation par objet, le fond est en gris. Au milieu, élimination des pixels situés en arrière de la source. En bas, découpage en zones avec le même nombre de sommets au dessus de l'horizon.*

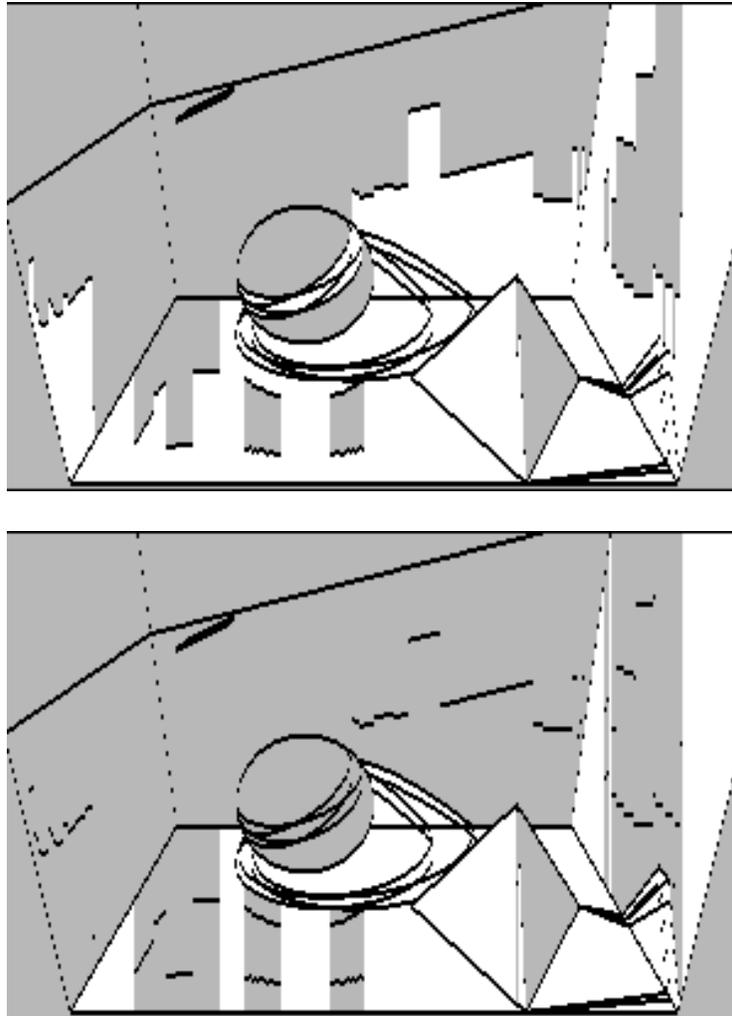


FIG. 7.16 - *Etapes de la segmentation (suite). En haut, découpage le long des discontinuités VE. Les segments de type ILLUM pour lesquels aucun obstacle potentiel n'a été trouvé sont rajoutés en gris. En bas, ajout des discontinuités E^* (visibles en haut de l'ombre de la pyramide à droite) et, en gris, extension des segments ILLUM à leurs voisins.*

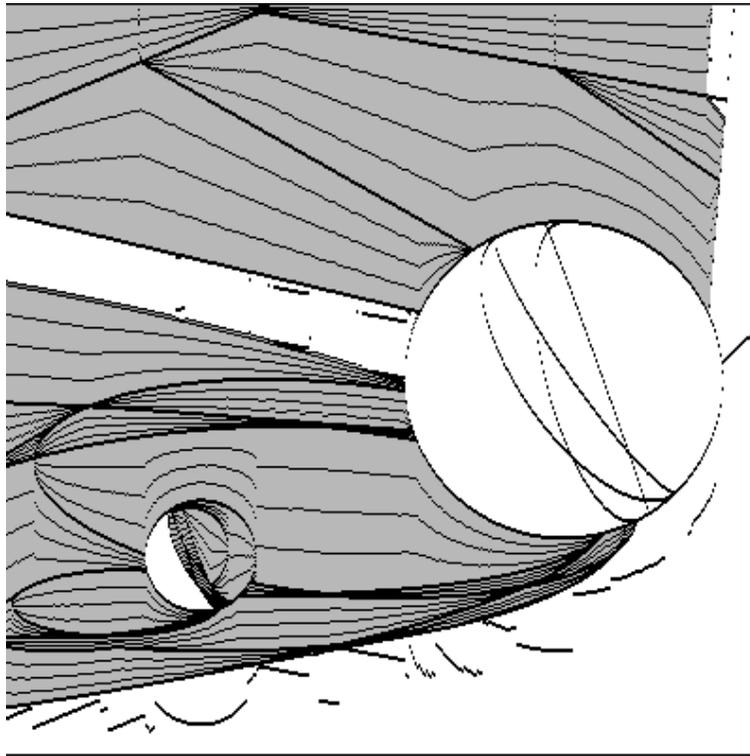


FIG. 7.17 - *Calcul de l'image de test par l'algorithme proposé. En blanc, les pixels pour lesquels aucun rayon d'ombrage n'a été lancé. En gris, les pixels pour lesquels des rayons ont été lancés vers les sommets de la source. En noir, les lignes monotones de pixels pour l'interpolation horizontale.*

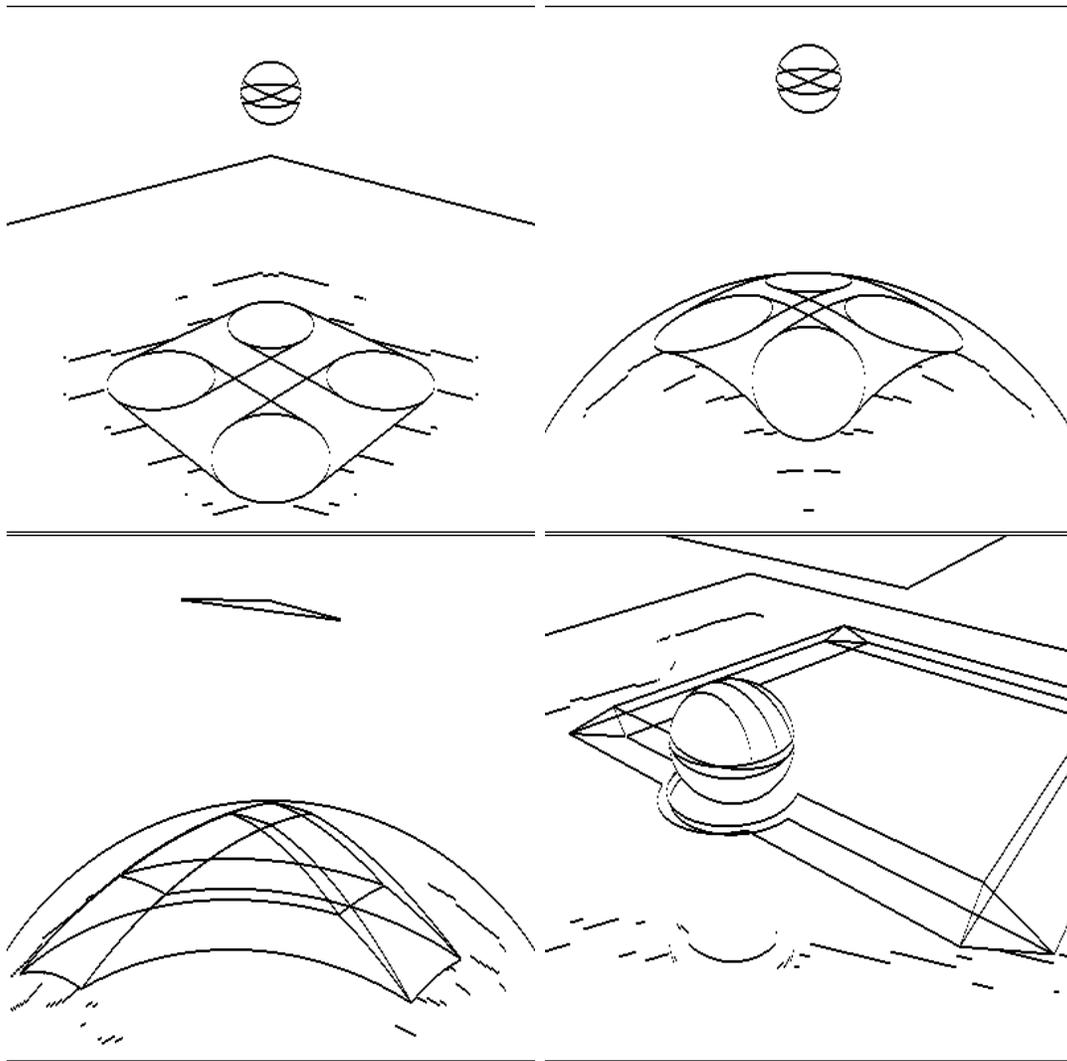


FIG. 7.18 - Exemples de découpages suivant les discontinuités.

Conclusion

Nous avons abordé, dans cette thèse, le problème du calcul de l'illumination d'une scène par des sources étendues. Le chapitre 1 rappelle la construction d'un modèle d'illumination globale à partir des caractéristiques locales de la réflexion de la lumière sur les surfaces. Cette construction est appliquée en particulier à la méthode de la radiosité qui s'obtient en faisant l'hypothèse lambertienne de la réflexion de la lumière et qui est le cadre de ce travail. Les différents algorithmes qui ont été développés pour calculer et afficher la distribution de lumière sont présentés. Nous introduisons la notion de vecteur de forme, obtenue en factorisant, dans l'expression du facteur de forme, la normale sur le récepteur. La radiosité scalaire est ainsi remplacée par un vecteur de radiosité. Nous montrons au chapitre 4 que l'emploi de ce vecteur permet de réduire la densité d'échantillonnage sur les surfaces gauches pour une même précision d'éclairage. La suite de la thèse se divise en deux parties : d'une part, le calcul et la représentation de l'illumination dans l'espace objet par l'intermédiaire d'un maillage des surfaces et d'autre part des méthodes d'accélération du calcul pixel par pixel de l'ombrage direct des lampes.

Le deuxième chapitre passe en revue les méthodes de maillage déjà proposées et en analyse les avantages et les défauts. Cette étude fournit les principes de la nouvelle méthode proposée qui sont de ne pas utiliser d'approximation polygonale des surfaces gauches, de s'affranchir de certaines contraintes topologiques sur le maillage et de permettre l'utilisation d'informations même approximatives sur les dérivées spatiales de la radiosité.

Le chapitre 3 donne un algorithme de calcul des dérivées fondé sur la méthode d'estimation du facteur de forme par lancer de rayons. L'emploi direct de différences finies entre échantillons voisins est rendu impossible par l'importante erreur périodique introduite par le test discret de la visibilité de la source. Nous proposons une méthode à faible coût ajouté qui estime le déplacement d'un bord d'obstacle virtuel sur la source pour un petit déplacement autour du point récepteur.

Ces dérivées et le vecteur de radiosité sont utilisés par l'algorithme de maillage décrit au chapitre 4. Le placement des échantillons, le raffinement du maillage et la reconstruction d'une fonction de radiosité continue sont vus en détail. Un choix judicieux des fonctions utilisées pour la reconstruction permet d'éviter les

contraintes de restriction des arbres quaternaires utilisés. Dans le chapitre 5, nous présentons des tests sur la solution de radiosité obtenue grâce à l'algorithme de maillage développé au chapitre précédent. Les composantes directe et indirecte de l'éclairage sont traitées séparément. Les résultats obtenus donnent des informations importantes sur le comportement de la solution de radiosité. En particulier, l'emplacement des erreurs de la composante indirecte et le peu d'influence de la finesse du maillage de l'éclairage direct sur cette composante permettent d'envisager la construction d'un maillage fixe pour l'éclairage indirect. La méthode de test proposée pourrait également être utilisée avec profit pour, par exemple, étudier le lien entre les paramètres de contrôle des méthodes de résolution hiérarchiques et la précision des résultats obtenus.

La méthode d'accélération du calcul de l'ombrage pour une source étendue du chapitre 6 utilise la transformation de l'intégrale de surface du facteur de forme en intégrale de contour. Un algorithme de *marching square* construit automatiquement une approximation polygonale du contour de la source, étendant ainsi cette technique d'intégration à des objets de forme quelconque. L'implémentation réalisée se compare favorablement aux techniques utilisant la méthode de Monte Carlo.

Finalement, le chapitre 7 étend la classification des discontinuités de l'illumination, qui n'existait jusque là que pour les scènes polyédriques, aux objets constitués de surfaces gauches. On montre ainsi que toutes les discontinuités séparant les zones de pleine illumination des pénombres sont calculables simplement. Un algorithme de segmentation des colonnes de pixels dans le plan image permet de réduire encore les calculs nécessaires à l'ombrage par des sources étendues en concentrant l'essentiel du travail dans les ombres et les pénombres.

Les différents algorithmes proposés dans cette thèse permettent de obtenir un système de production d'images de qualité pour la radiosité. La composante indirecte peut être calculée sur un maillage fixe mais adapté par un traitement préliminaire. En effet, les résultats du chapitre 5 suggèrent que les zones où l'erreur est la plus importante sont plus liées à la géométrie locale de la scène (coins, fuites d'ombre, etc.) qu'au placement des sources primaires et que la précision de calcul de la composante directe a une influence réduite.

La composante directe est recalculée avec la précision spatiale adaptée au moment du calcul de l'image en utilisant les techniques présentées dans les deux derniers chapitres. Il faut noter que l'algorithme de tracé des discontinuités n'est pas lié à l'utilisation du lancer de rayons pour obtenir une image de la scène mais peut aussi être appliqué à une grille bidimensionnelle de points sur une surface pour produire une texture d'illumination utilisable pour produire ensuite plusieurs images par une méthode plus simple comme le z-buffer. Cependant plusieurs points restent à développer en particulier dans le traitement d'objets plus complexes (calcul des secteurs plans de discontinuité) et dans les méthodes d'interpolation utilisées à l'intérieur des pénombres.

Table des matières

1	Modèles d'illumination et radiosité	5
1.1	Réflexions et interréflexions de la lumière	5
1.2	Algorithmes pour la radiosité	10
1.2.1	Calcul des facteurs de forme	10
1.2.2	Résolution du système d'équilibre global	14
1.2.3	Production d'images	15
1.3	Vecteur de forme	16
2	Méthodes de maillage pour la radiosité	19
2.1	Maillages polygonaux adaptatifs	20
2.1.1	Maillage à plusieurs niveaux de détail	20
2.1.2	Maillage cohérent	21
2.2	Maillages utilisant des arbres binaires de partition de l'espace	21
2.3	Tracé des discontinuités	22
2.4	Textures de radiosité	25
2.5	Échantillons d'illumination en 3D	26
2.6	Bases de fonctions globales	26
2.7	Méthodes hiérarchiques	27
2.8	Bilan	28
3	Dérivées premières de l'illumination	31
3.1	Utilisation du vecteur de forme	31
3.2	Origine des variations d'illumination	32
3.3	Calcul de la dérivée	33
3.3.1	Variation d'angle solide	33
3.3.2	Variation de la partie visible	34
3.4	Résultats	37
4	Maillage adaptatif sans polygones	43
4.1	Échantillonnage de l'illumination	44
4.1.1	Placement des échantillons	44
4.1.2	Critères pour le raffinement	47
4.2	Reconstruction de la radiosité	49

4.2.1	Filtres elliptiques	50
4.2.2	Filtres circulaires et extrapolation	50
4.2.3	Fonctions de pondération et subdivision	51
4.3	Composante directe et indirecte de l'illumination	54
4.4	Radiosité et surfaces courbes	55
4.5	Conclusion	56
5	Tests de la radiosité reconstruite	61
5.1	Méthodes de test	61
5.2	Eclairage direct	63
5.2.1	Calcul de la référence	63
5.2.2	Subdivision adaptative	64
5.2.3	Dérivées de l'illumination	64
5.3	Eclairage indirect	65
5.3.1	Notion de cohérence	65
5.3.2	Fonction reconstruite et fonction intégrée	68
5.3.3	Fuites d'ombre et maillage propre à l'éclairage indirect	70
5.3.4	Influence du maillage de l'éclairage direct	70
5.4	Conclusion	71
6	Intégration de contour approché	81
6.1	Motivations	81
6.2	Construction du contour approché	84
6.2.1	Principe	84
6.2.2	Marching squares	84
6.3	Raffinement du contour approché	85
6.3.1	Critère	86
6.3.2	Problèmes avec le raffinement	87
6.4	Remarques sur l'implémentation	89
6.4.1	Structures de données	89
6.4.2	Alternative à la dichotomie	89
6.5	Calcul des dérivées	92
6.6	Résultats	93
6.6.1	Algorithme de Monte Carlo adaptatif	93
6.6.2	Méthode de calcul des images	93
6.6.3	Comparaison des deux méthodes	94
6.7	Conclusion	96
7	Maillage discret des discontinuités pour des surfaces courbes	99
7.1	Discontinuités de l'illumination sur des surfaces courbes	99
7.1.1	Introduction	99
7.1.2	Classification	101
7.1.3	Principe de l'algorithme	103

7.2	Algorithme de calcul d'image	105
7.2.1	Calcul et tri des discontinuités E^*	105
7.2.2	Détection des discontinuités	108
7.2.3	Segmentation des colonnes de pixels	109
7.2.4	Groupement de segments en bandes	111
7.2.5	Interpolation de l'illumination	113
7.2.6	Candidats pour le lancer de rayons	115
7.3	Résultats	116

Table des figures

1.1	Géométrie pour la luminance.	6
1.2	Géométrie pour la réflexion.	7
1.3	Géométrie pour les facteurs de forme.	11
1.4	Équivalent de Nusselt pour le calcul des facteurs de forme.	12
1.5	Géométrie pour l'intégration de contour polygonal.	14
1.6	Calcul de l'énergie incidente sur un pixel pour un modèle de prise de vue simple.	15
2.1	Discontinuités pour une source triangulaire et un obstacle rectan- gulaire.	23
2.2	Discontinuité de type EEE.	24
2.3	Comparaison entre méthodes de maillage.	27
3.1	Géométrie de A_j^M et $A_j^{M'/M}$	33
3.2	Couple de cellules (C_a, C_b) pour le parcours suivant les lignes . . .	35
3.3	Géométrie de l'obstacle virtuel.	36
3.4	Géométrie pour les courbes.	37
3.5	Courbe de référence du facteur de forme calculée par intégration de contour.	39
3.6	Courbe de référence de la dérivée partielle par rapport à x du facteur de forme calculée par intégration de contour.	39
3.7	Courbe du facteur de forme calculée par la méthode de Wallace avec 64 échantillons par point.	40
3.8	Courbe de la dérivée partielle par rapport à x du facteur de forme calculée par différences finies.	40
3.9	Courbe de la dérivée partielle par rapport à x du facteur de forme calculée par la méthode proposée.	41
3.10	Erreur relative entre la méthode proposée et la référence.	41
4.1	Placement des échantillons.	45
4.2	Adaptation de la subdivision d'une surface paramétrique.	46
4.3	Distribution régulière dans l'espace paramétrique (à gauche) et dans l'espace objet (à droite) pour la sphère.	47

4.4	Point choisi pour imposer la coïncidence entre la fonction w_i de P_i et la somme des fonctions w_i^k des sous-échantillons P_i^k	53
4.5	Fuite de lumière et fuite d'ombre.	55
4.6	Fonctions gaussiennes pour un échantillon.	58
4.7	Défaut de reconstruction dû à un trop grand écart de niveau entre échantillons voisins.	59
4.8	Utilisation du vecteur de forme pour reconstruire l'illumination sur les surfaces courbes.	60
5.1	Scène utilisée pour les tests de cohérence.	62
5.2	Calcul de l'éclairage indirect par intégration de la lumière incidente sur l'hémisphère.	67
5.3	Vue de la seconde scène montrant la composante indirecte de l'éclairage seule.	68
5.4	Images de la scène deux calculées à partir des maillages β et $\beta + 3$	73
5.5	Exemples de maillages de la scène 2	74
5.6	Profil d'éclairage dans la traversée d'une pénombre. Maillage régulier et maillages adaptatifs de profondeurs maximales croissantes.	74
5.7	Utilisation des dérivées pour la reconstruction de l'éclairage direct.	75
5.8	Comparaison entre les radiosités reconstruites avec et sans dérivées	76
5.9	Profils d'éclairage indirect pour la «boite de Cornell».	77
5.10	Profils d'éclairage indirect pour la seconde scène.	78
5.11	Élimination des fuites d'ombre.	79
5.12	Profils d'éclairage le long de la ligne c pour la seconde scène.	79
5.13	Histogrammes de l'écart entre valeur reconstruite et valeur intégrée pour les pixels de l'image 128×128	80
6.1	Effet de marches d'escalier dû à l'approximation de la source étendue par un nombre trop faible de sources ponctuelles (16 dans ce cas).	82
6.2	Comparaison entre subdivision adaptative de la surface et approximation de son contour.	83
6.3	Configurations de passage du contour.	84
6.4	Variation périodique de la surface de source négligée.	85
6.5	Sommets ajoutés au contour pour tester sa validité.	86
6.6	Incohérences créées par le raffinement d'un rectangle.	88
6.7	Rayon limite pour un polygone.	90
6.8	Rayon limite pour une sphère.	91
6.9	Schéma de la scène de test vue de dessus.	95
6.10	Image de référence.	96
6.11	Détail des quatre images. Images 1, 3, 4 et 5, de haut en bas et de gauche à droite.	98

7.1	Les singularités de codimension 0 pour des objets lisses par morceaux.	100
7.2	Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types EE, EF et FF).	101
7.3	Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types EC et FC).	102
7.4	Les singularités multilocales de codimension 1 pour des objets lisses par morceaux (types EEE, EEF, EFF et FFF).	103
7.5	Passages d'une singularité multilocale de type point triple.	104
7.6	Secteurs angulaires de discontinuité VE.	106
7.7	Tri des plans supports des discontinuités E^* .	107
7.8	Rayon IS utilisé pour vérifier l'existence réelle de la discontinuité EV.	108
7.9	Deux types différents de segment d'où tous les sommets de la source sont visibles.	111
7.10	Organisation des colonnes de pixels et de segments pour permettre le rattachement en bandes de segments adjacents.	112
7.11	Lignes d'échantillons pour l'interpolation horizontale.	113
7.12	Algorithme d'interpolation horizontale.	114
7.13	Volumes de pénombre avec les plans de discontinuité.	115
7.14	Propagation des discontinuités maximales dans la hiérarchie des boîtes englobantes.	118
7.15	Étapes de la segmentation.	119
7.16	Étapes de la segmentation (suite).	120
7.17	Calcul de l'image de test par l'algorithme proposé.	121
7.18	Exemples de découpages suivant les discontinuités.	122

Bibliographie

- [AWG78] Peter Atherton, Kevin Weiler, and Donald Greenberg. Polygon shadow generation. *Computer Graphics (SIGGRAPH '78 Proceedings)*, 12(3):275–281, Aug. 1978.
- [BB84] L. Brotman and Norman Badler. Generating soft shadows with a depth buffer algorithm. *IEEE Computer Graphics and Applications*, 4(10):71–81, Oct. 1984.
- [BMSW91] Daniel R. Baum, Stephen Mann, Kevin P. Smith, and James M. Winget. Making radiosity usable: Automatic preprocessing and meshing techniques for the generation of accurate radiosity solutions. *Computer Graphics (SIGGRAPH '91 Proceedings)*, 25(4):51–60, July 1991.
- [BRW89] Daniel R. Baum, Holly E. Rushmeier, and James M. Winget. Improving radiosity solutions through the use of analytically determined form-factors. *Computer Graphics (SIGGRAPH '89 Proceedings)*, 23(3):325–34, July 1989.
- [Cam91] Alvin T. Campbell, III. *Modeling Global Diffuse Illumination for Image Synthesis*. PhD thesis, Dept. of Computer Sciences, Univ. of Texas at Austin, December 1991.
- [CCWG88] Michael Cohen, Shenchang Eric Chen, John R. Wallace, and Donald P. Greenberg. A progressive refinement approach to fast radiosity image generation. *Computer Graphics (SIGGRAPH '88 Proceedings)*, 22(4):75–84, Aug. 1988.
- [CF89] Norman Chin and Steven Feiner. Near real-time shadow generation using bsp trees. *Computer Graphics (SIGGRAPH '89 Proceedings)*, 23(3):99–106, July 1989.
- [CF90] Alvin T. Campbell, III and Donald S. Fussell. Adaptive mesh generation for global diffuse illumination. *Computer Graphics (SIGGRAPH '90 Proceedings)*, 24(4):155–64, Aug. 1990.

- [CF92] Norman Chin and Steven Feiner. Fast object-precision shadow generation for area light sources using bsp trees. In *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*, pages 21–30, March 1992.
- [CG85] Michael Cohen and Donald P. Greenberg. The hemi-cube: A radiosity solution for complex environments. *Computer Graphics (SIGGRAPH '85 Proceedings)*, 19(3):31–40, Aug. 1985. also in Tutorial: Computer Graphics: Image Synthesis, Computer Society Press, Washington, 1988.
- [CGIB86] Michael Cohen, Donald P. Greenberg, Dave S. Immel, and Philip J. Brock. An efficient radiosity approach for realistic image synthesis. *IEEE Computer Graphics and Applications*, 6(3):26–35, March 1986.
- [CPC84] Robert L. Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. *Computer Graphics (SIGGRAPH '84 Proceedings)*, 18(3):137–45, July 1984.
- [Cro77] Franklin Crow. Shadow algorithms for computer graphics. *Computer Graphics (SIGGRAPH '77 Proceedings)*, 11(3):242–248, Aug. 1977.
- [CT82] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics*, 1(1):7–24, January 1982.
- [Dev88] Olivier Devillers. Méthodes d'optimisation du lancer de rayon. Technical Report Thèse, Université de Paris Sud, Centre d'Orsay, Jun.. 1988.
- [DW85] Mark A. Z. Dippe and Erling Henry Wold. Antialiasing through stochastic sampling. *Computer Graphics (SIGGRAPH '85 Proceedings)*, 19(3):69–78, July 1985.
- [EC90] Gershon Elber and Elaine Cohen. Hidden curve removal for free form surfaces. *Computer Graphics (SIGGRAPH '90 Proceedings)*, 24(4):95–104, August 1990.
- [Fra82] Richard Franke. Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38(157):181–200, Jan. 1982.
- [GM90] Ziv Gigus and Jitendra Malik. Computing the aspect graph for line drawings of polyhedral objects. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(2):113–122, Feb. 1990.
- [GS87] Jeffrey Goldsmith and John Salmon. *Automatic Creation of Object Hierarchies for Ray Tracing*, volume 7. IEEE Computer Graphics and Applications, May 1987.

- [GTGB84] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modelling the interaction of light between diffuse surfaces. *Computer Graphics (SIGGRAPH '84 Proceedings)*, 18(3):212–22, July 1984.
- [Han92] Pat Hanrahan. From radiometry to the rendering equation. In *SIGGRAPH '92 Radiosity course notes*, volume 11. Aug. 1992.
- [Hec90] Paul Heckbert. Adaptive radiosity textures for bidirectional ray tracing. *Computer Graphics (SIGGRAPH '90 Proceedings)*, 24(4):145–54, Aug. 1990.
- [Hec91] Paul Heckbert. *Simulating Global Illumination Using Adaptive Meshing*. PhD thesis, CS Division (EECS), Univ. of California, Berkeley, June 1991.
- [HS90] Pat Hanrahan and David Salzman. A rapid hierarchical radiosity algorithm for unoccluded environments. *Proceedings Eurographics Workshop on Photosimulation, Realism and Physics in Computer Graphics*, pages 151–71, June 1990.
- [HSA91] Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. *Computer Graphics (SIGGRAPH '91 Proceedings)*, 25(4):197–206, July 1991.
- [HW91] Eric A. Haines and John R. Wallace. Shaft culling for efficient ray-traced radiosity. *Second Eurographics Workshop on Rendering*, May 1991.
- [Kaj86] James T. Kajiya. The rendering equation. *Computer Graphics (SIGGRAPH '86 Proceedings)*, 20(4):143–50, Aug. 1986.
- [KJ92] Arjan Kok and Frederik Jansen. Adaptive sampling of area light sources in ray tracing including diffuse interreflection. *Computer Graphics Forum (Eurographics '92 Proceedings)*, 2(3):289–298, Sept. 1992.
- [LD90] Charles Loop and Tony DeRose. Generalized b-spline surfaces of arbitrary topology. volume 24, pages 347–356, August 1990.
- [LRU85] Mark E. Lee, Richard A. Redner, and Samuel P. Uselton. Statistically optimized sampling for distributed ray tracing. *Computer Graphics (SIGGRAPH '85 Proceedings)*, 19(3):61–7, July 1985.
- [LTG92] Dani Lichinski, Filippo Tampieri, and Donald P. Greenberg. Discontinuity meshing for accurate radiosity. *IEEE Computer Graphics and Applications*, 12(6):25–39, November 1992.

- [Mal88] Thomas J.V. Malley. A shading method for computer generated images. Technical Report Master's Thesis, June 1988.
- [MN88] Don P. Mitchell and Arun N. Netravali. Reconstruction filters in computer graphics. *Computer Graphics (SIGGRAPH '88 Proceedings)*, 22(4):221,228, August 1988.
- [NN83] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of 3d-objects taking account of shadows and interreflection. In *COMPSAC '83 Proc. IEEE 7th Intl. Comp. Soft. and Applications Conf.*, pages 237–242, November 1983.
- [NN85] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of three-dimensional objects taking account of shadows and interreflection. *Computer Graphics (SIGGRAPH '85 Proceedings)*, 19(3):23–30, July 1985.
- [PK89] Jean Ponce and David J. Kriegman. An algebraic approach to computer vision. *Revue d'intelligence artificielle*, 3(2):105–136, 1989.
- [PSV90] Claude Puech, François Sillion, and Christophe Vedel. Improving interaction with radiosity-based lighting simulation programs. *Computer Graphics (1990 Symposium on Interactive 3D Graphics)*, 24(2):51–7, March 1990.
- [Rie87] J.H. Rieger. On the classification of views of piecewise-smooth objects. *Image and Vision Computing*, 5:91–97, 1987.
- [RPV93] Holly Rushmeier, Charles Patterson, and Aravindan Veerasamy. Geometric simplification for indirect illumination calculations. In *Proceedings of Graphics Interface '93*, pages 227–236, Toronto, Ontario, May 1993.
- [SAS92] Brian E. Smits, James R. Arvo, and David H. Salesin. An importance-driven radiosity algorithm. *Computer Graphics (SIGGRAPH '92 Proceedings)*, 26(4):273–282, July 1992.
- [She68] Donald Shepard. A two-dimensional interpolation function for irregularly spaced data. In *23rd Nat. Conf. ACM*, pages 517–523, 1968.
- [Shi90] Peter Shirley. Physically based lighting calculations for computer graphics. Technical Report PhD thesis, 1990.
- [SP89] François Sillion and Claude Puech. A general two-pass method integrating specular and diffuse reflection. *Computer Graphics (SIGGRAPH '89 Proceedings)*, 23(3):335–44, July 1989.

- [Ved92] Christophe Vedel. Improved storage and reconstruction of light intensities on surfaces. In *Third Eurographics Workshop on Rendering*, Bristol, England, May 1992.
- [Ved93] Christophe Vedel. Computing illumination from area light sources by approximate contour integration. *Proceedings of Graphics Interface '93*, pages 237–244, May 1993.
- [WEH89] John R. Wallace, Kells A. Elmquist, and Eric A. Haines. A ray tracing algorithm for progressive radiosity. *Computer Graphics (SIGGRAPH '89 Proceedings)*, 23(3):315–24, July 1989.
- [WH92] Gregory J. Ward and Paul Heckbert. Irradiance gradients. In *Third Eurographics Workshop on Rendering*, Bristol, England, May 1992.
- [WPF90] Andrew Woo, Pierre Poulin, and Alain Fournier. A survey of shadow algorithms. *IEEE Computer Graphics and Applications*, 10(6):13–32, Nov. 1990.
- [WRC88] Gregory J. Ward, Francis M. Rubinstein, and Robert D. Clear. A ray tracing solution for diffuse interreflection. *Computer Graphics (SIGGRAPH '88 Proceedings)*, 22(4):85–92, Aug. 1988.
- [Zat92] Harold R. Zatz. Galerkin radiosity: a higher order solution method for global illumination. Technical Report Master's Thesis, Program of Computer Graphics, Cornell Univ., Aug. 1992.

Mots-clefs : radiosité, maillage, gradients d'illumination, reconstruction, intégration de contour, discontinuités, surfaces gauches.

Résumé : Cette thèse concerne la simulation de l'éclairage par des sources de lumière étendues dans le cadre de la méthode dite de radiosité. Le problème est abordé sous deux angles distincts : calcul dans l'espace objet par l'intermédiaire d'un maillage et calcul dans l'espace image pixel par pixel. Les deux premiers chapitres sont consacrés à la méthode de radiosité et aux algorithmes existants qui sont le point de départ de ce travail. Les méthodes de maillage utilisées pour représenter la fonction d'intensité lumineuse sur les surfaces d'une scène sont plus particulièrement présentées. Les trois chapitres suivants introduisent une méthode de maillage qui n'est pas limitée à des approximations polyédriques des environnements et qui utilise les dérivées de la fonction d'illumination pour améliorer sa reconstruction à partir des échantillons. Deux nouvelles méthodes d'accélération du calcul de l'éclairage pour des sources polygonales sont décrites dans les derniers chapitres : l'intégration de contour approché et le tracé des discontinuités de l'illumination dans l'espace image. Toutes deux sont des extensions aux surfaces gauches de méthodes existantes qui étaient limitées aux polyèdres.

Keywords: radiosity, meshing, illumination gradients, reconstruction, contour integration, discontinuities, curved surfaces.

Abstract: This thesis deals with the simulation of lighting by extended light sources in the framework of the radiosity method. The problem is treated with two distinct approaches: computation in object space through the use of a mesh and computation in image space pixel by pixel. The first two chapters describe the radiosity method and existing algorithms that are the starting point of this work. In particular, the meshing techniques used to code the light intensity function on the surfaces of a scene are reviewed. The following three chapters introduce a meshing method not limited to polyhedral approximation of the environment and using the derivatives of the illumination function to improve its reconstruction from the samples. Two new acceleration techniques for the computation of lighting from polygonal sources are described in the last chapters: approximate contour integration and discontinuity tracing in image space. Both are extensions to curved surfaces of previous methods that were only valid for polyhedra.