

Le rendu par lancer de rayons (ray tracing)

PRINCIPE

Algorithme de rendu pour l'obtention d'images de qualité photo-réaliste

ALGORITHME

AVANTAGES

- Diffusions
- Réflexions spéculaires
- Transmissions
- Ombres portées
- Elimination des parties cachées
- ...

INCONVÉNIENTS

EXEMPLES

MATHÉMATIQUES

OPTIMISATIONS

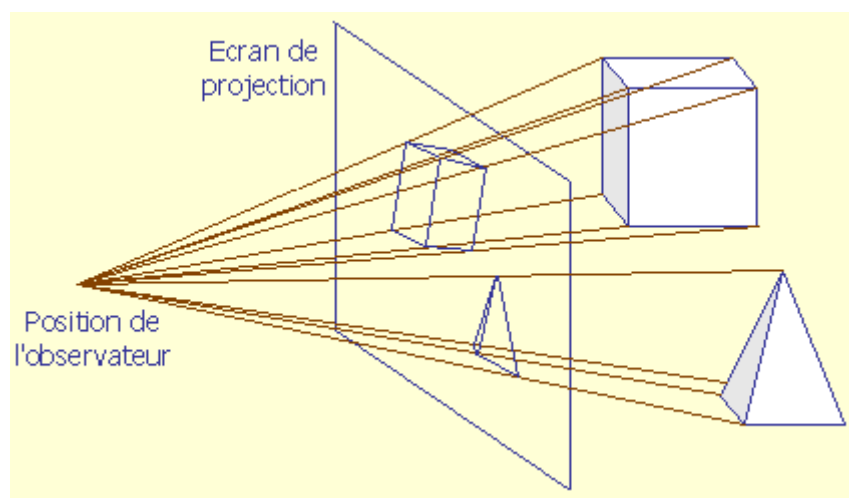
CONCLUSION



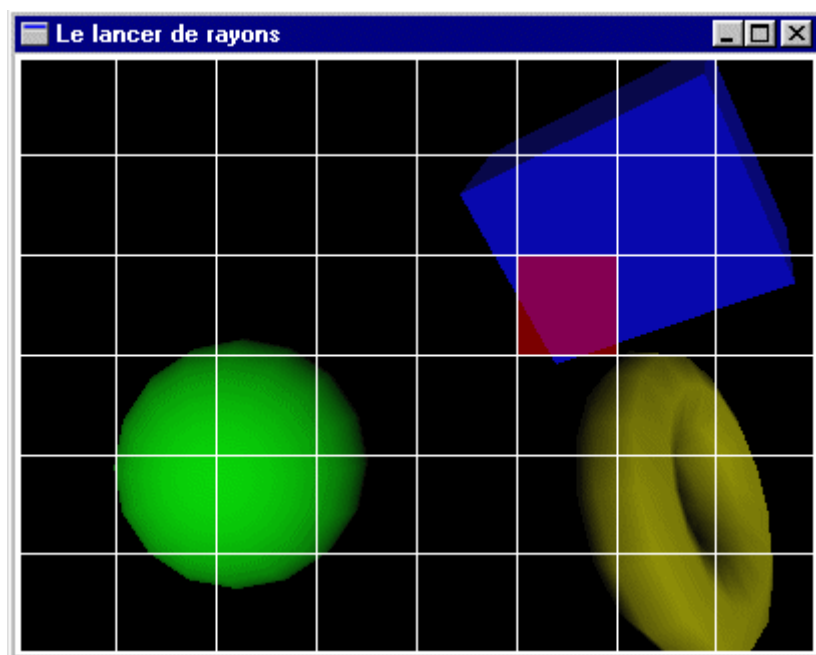
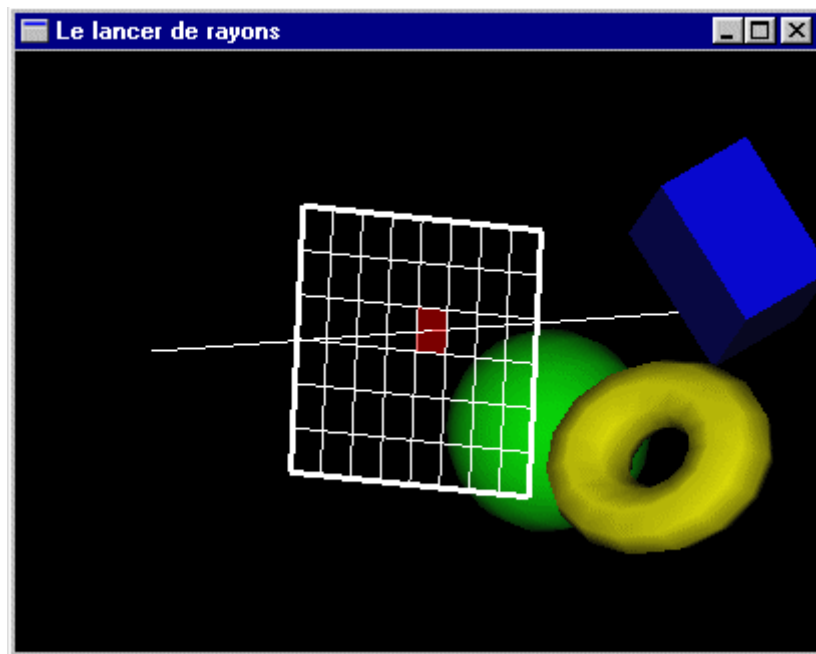
Principe

Pour déterminer comment ils ont été créés, on suit en sens inverse le parcours des rayons lumineux qui s'impriment sur la rétine de l'observateur.

En partant depuis l'œil de l'observateur, on traverse l'écran virtuel sur lequel est matérialisée l'image créée, puis on "entre" dans la scène de manière à déterminer de quel(s) objet(s) un rayon lumineux provient, et ainsi, à déterminer ses caractéristiques chromatiques.



On traite le rayon primaire partant de l'observateur et passant par chaque pixel de l'écran virtuel de manière à déterminer la couleur attribuée à ce pixel.

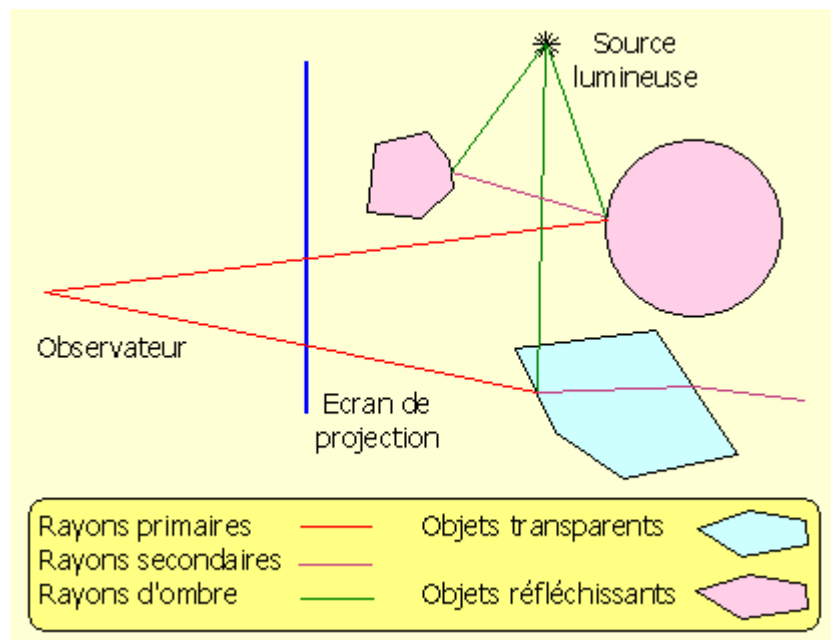


[Le programme Aux](#)

[Le programme GLUT](#)

[Exécutable Aux](#)

[Exécutable GLUT](#)



Rayons secondaires
_ □ ×

<u>Le programme Aux</u>	<u>Le programme GLU+</u>
<u>Exécutable Aux</u>	<u>Exécutable GLU+</u>

Lors du parcours d'un rayon lumineux dans une scène, deux cas sont possibles:

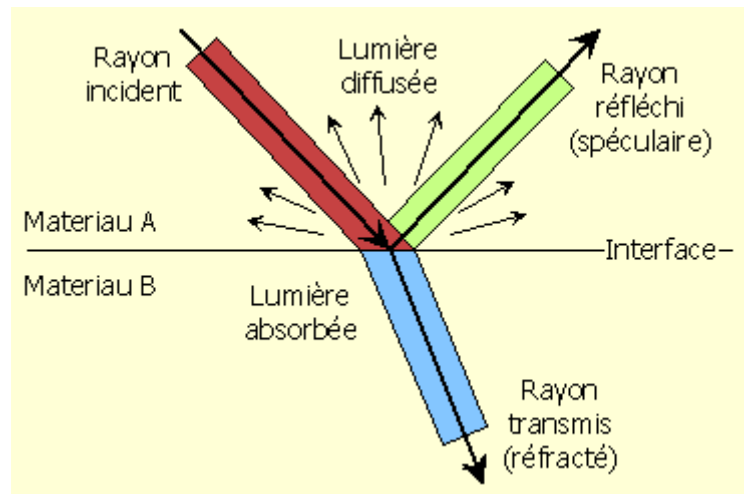
(1) Aucun objet n'intercepte le rayon.

-> La couleur est celle du fond (noir ou autre).

(2) Un ou plusieurs objets interceptent le rayon.

- On trouve l'objet le plus proche de la source du rayon (l'observateur pour les rayons primaires, une intersection rayon-objet pour les rayons secondaires).

- En fonction des caractéristiques de cet objet vis à vis de la lumière, la teinte du pixel est calculée par somme des composantes de lumière (a) diffusée, (b) réfléchie et (c) transmise par l'objet au point d'intersection (composantes qui ont "créé" ce rayon "incident").



Décomposition au niveau d'une interface
_ □ ×

<u>Le programme Aux</u>	<u>Le programme GLUT</u>
<u>Exécutable Aux</u>	<u>Exécutable GLUT</u>

(a) Valeur diffusée:

On doit:

- Déterminer le point d'intersection P entre l'objet et le rayon incident
- Lancer un rayon de P vers chacune des sources lumineuses présentes dans la scène pour déterminer si ce rayon est intercepté ou non par un objet (à l'ombre ou non)

- Pour chaque rayon non intercepté, appliquer une formule adaptée au calcul de la lumière diffuse (loi de Lambert ou autre) avec utilisation des coefficients de diffusion de la surface de l'objet
- Sommer le résultat de chacun de ces calculs pour obtenir la quantité globale de lumière diffusée

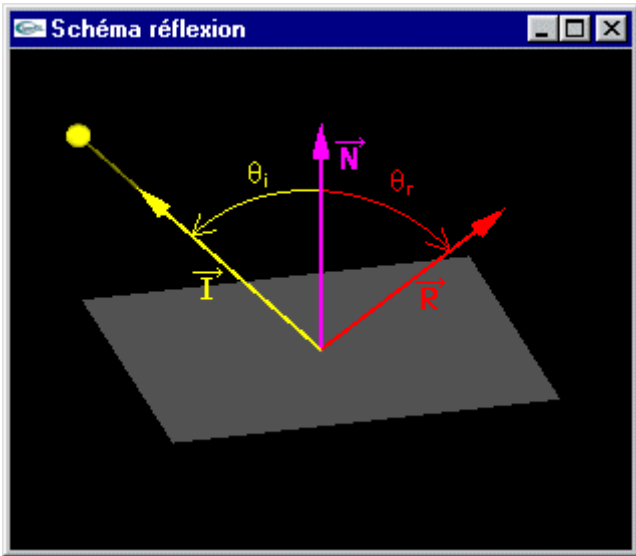
(b) La valeur obtenue par réflexion est calculée en lançant récursivement l'algorithme de lancer de rayons sur le rayon secondaire défini avec pour origine le point P et pour direction la direction de réflexion du rayon initial puis en pondérant la valeur obtenue par les coefficients de réflexion de la surface de l'objet.

Les rayons incident, de réflexion et la normale à l'interface au point d'incidence sont dans le même plan.

L'angle de réflexion θ_r est tel que $\theta_r = -\theta_i$ où θ_i est "l'angle d'incidence" par rapport à la normale à l'interface au niveau du point de réflexion.

Angle de réflexion

Schéma réflexion



Valeurs

INCIDENT	:	-0.684	0.669	-0.289
NORMALE	:	0.000	1.000	0.000
REFLEXION	:	0.684	0.669	0.289
THETA I	:	47.988 degrees		
THETA R	:	47.988 degrees		

Le programme GLUT

Exécutable GLUT

Les rayons secondaires de réflexion sont exclusivement tracés dans le milieu inter-objets.

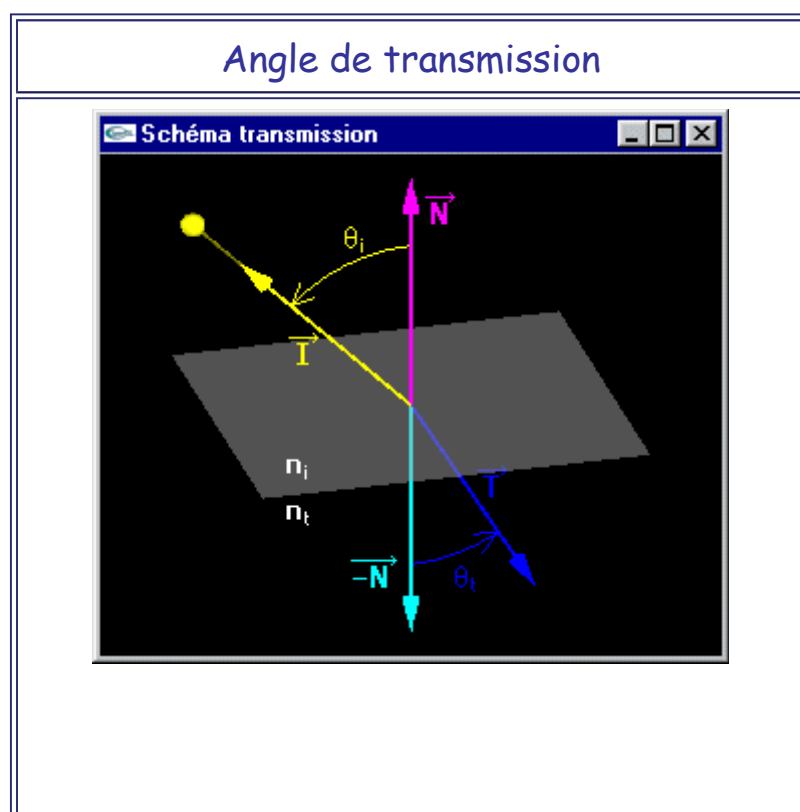
(c) La valeur obtenue par transmission est calculée en lançant récursivement l'algorithme de lancer de rayons sur le rayon secondaire défini avec pour origine le point P et pour direction la direction de transmission puis en pondérant cette valeur par les coefficients de transmission de la surface de l'objet.

Lors d'une réfraction, la direction du rayon incident est modifiée selon la loi de Snell:

$$n_i \sin(\theta_i) = n_t \sin(\theta_t)$$

où n_i est l'indice de réfraction du matériau dans lequel se propage le rayon incident, n_t est l'indice de réfraction du matériau dans lequel se propage le rayon transmis, θ_i est "l'angle d'incidence" par rapport à la normale \vec{N} à l'interface orientée dans le milieu incident et calculée au niveau du point de transmission et θ_t est "l'angle de transmission" par rapport à $-\vec{N}$.

Les rayons incident, de transmission et la normale à l'interface au point d'incidence sont dans le même plan.



Valeurs			
n = ni/nt	:	0.727	
INCIDENT	:	-0.707	0.707 0.000
NORMALE	:	0.000	1.000 0.000
TRANSMISSION	:	0.514	-0.858 0.000
THETA I	:	45.000 degres	
THETA T	:	30.949 degres	

Le programme GLUt

Exécutable GLUt

Les rayons secondaires de transmission sont tracés dans le milieu inter-objets mais aussi dans les objets "transparents".
 -> Cette caractéristique contraint le programmeur à gérer les entrées/sorties des objets.

Remarque

En fonction de n_i , n_t et θ_i , la valeur de θ_t peut ne pas être définie (cas où le rayon incident fait un angle avec la normale suffisant pour ne pas être transmis et être totalement réfléchi (effet miroir)).

Algorithme simplifié

```

Programme lancerDeRayons
  scene s
  rayon r
  tableau_de_couleur teinte
  s <- scène à afficher
  pour chaque pixel p(x,y) de l'écran
    r <- rayon passant par l'observateur et p
    teinte[x,y] <- lancerRayon(s,r)
  finpour
finprogramme

couleur lancerRayon(scene s,rayon r)
  couleur cr,ct,cd
  rayon rr,rt,rd
  coefficient kr,kt
  point p
  objet o
  début
  o <- objet de la scène s dont l'intersection
  avec r est la plus proche de la source de r
  si o n'existe pas
    retourner (noir)

```

```

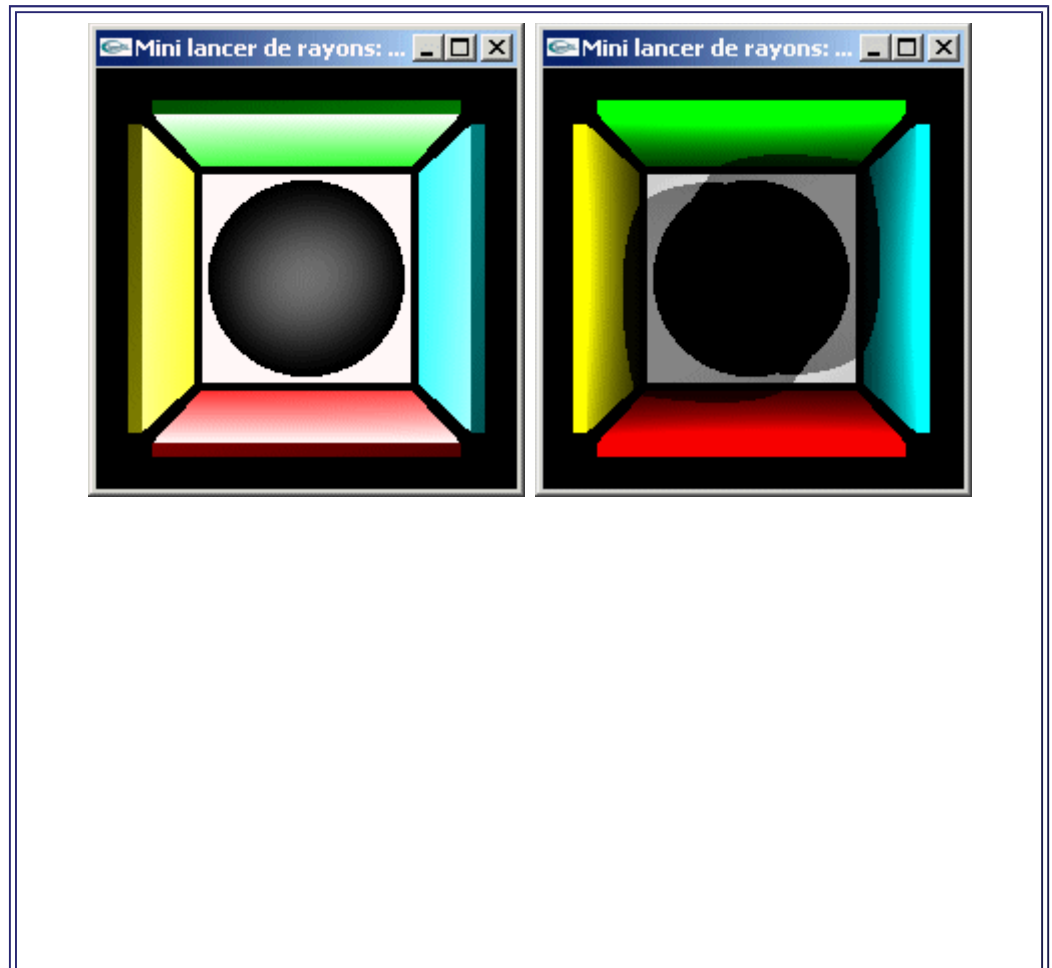
sinon
p <- point d'intersection entre r et o
rr <- rayon réfléchi de r en p sur o
rt <- rayon transmis de r en p vis à vis de o
kr <- coefficients de réflexion de o
kt <- coefficients de transmission de o
cr <- lancerRayon(s,rr) * kr
ct <- lancerRayon(s,rt) * kt
cd <- noir
pour chaque lumiere sl de s
  rd <- rayon de p vers sl
  si rd non intercepté par un objet
    cd <- cd + diffusion(sl,o,rd)
  finsi
finpour
retourner(cr + cd + ct)
finsi
finfonction

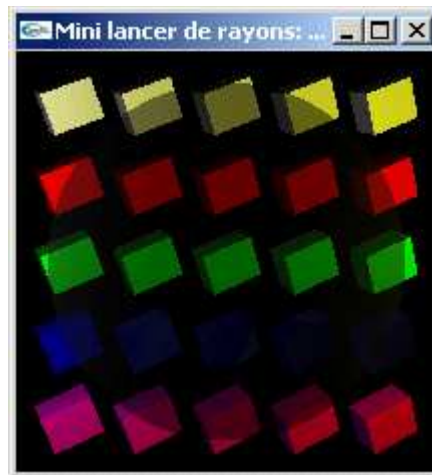
```

Avantages du lancer de rayons

Réalisme

Des effets tels que les réflexions et les transparences sont restitués intrinsèquement. Aucun autre algorithme n'est capable de modéliser ces effets de manière aussi satisfaisante.





Le programme GLUT

Exécutable GLUT

Les parties cachées sont éliminées de manière intrinsèque. Au niveau 1 de la récursivité, on est en présence d'un algorithme très voisin du Z-Buffer dans ses résultats: le Ray Casting. Les ombres portées sont intrinsèques au modèle.

Objets de base évolués

Pas de décomposition de la scène en facettes triangulaires ou surfaces élémentaires planes (contrairement au Z-Buffer)
-> Les objets courbes apparaissent réellement courbes.

Les opérations de base utilisées sont les tests d'intersection objet->rayon. Ces tests sont implantés au moyen des équations cartésiennes et paramétriques par résolution de systèmes d'équations.

Classiquement les objets modélisant une scène affichée en lancer de rayons sont des objets représentables par équation(s) mathématique(s): des sphères, des parallélépipèdes rectangles, des cônes, des surfaces contrôlées, ... (affectés de rotations, translations, zooms).

Bien entendu, on peut aussi utiliser des ensembles de facettes.

Inconvénients du lancer de rayons

Récurtivité

A chaque interception d'un objet par un rayon, outre les rayons d'ombre, on lance potentiellement deux rayons: un rayon pour la gestion de la transparence et un rayon pour la gestion de la réflexion spéculaire.

-> On peut assister à une explosion exponentielle du nombre total de rayons tracés à l'intérieur de la scène. Si n est le niveau de profondeur de la récursivité, il y a au maximum de l'ordre de 2^n rayons tracés pour chaque pixel.

Pour obtenir un certain réalisme le niveau de récursivité doit être poussé assez loin (au moins 8, habituellement 10 à 12 voire 14) et donc les temps de calcul peuvent être très longs (plusieurs heures par image pour des scènes complexes).

Problèmes de calcul d'illumination

L'illumination est "moyennement" locale.

Tout illumination provient d'une diffusion sous l'éclairage d'une source de lumière répertoriée comme telle. Cet éclairage est soit direct, soit le résultat d'un nombre de réflexions

spéculaires ou de transmissions successives égal au maximum au niveau de profondeur de récursivité utilisé.

Dans la réalité toutes les surfaces et tous les milieux sont récepteurs et émetteurs de lumière et les réflexions et transmissions sont infinies.

-> En lancer de rayon, il peut apparaitre des zones non éclairées qui auraient du l'être un peu.

-> En lancer de rayon, il peut apparaitre des zones illuminées très crument qui n'auraient peut-être pas du être illuminées aussi fortement.

Ce problème est partiellement résolu en attribuant une composante de lumière ambiante à chacun des éléments de la scène (composante, souvent constante car on ne sait pas la calculer précisément avec l'algorithme de ray-tracing, ajoutée aux autres).

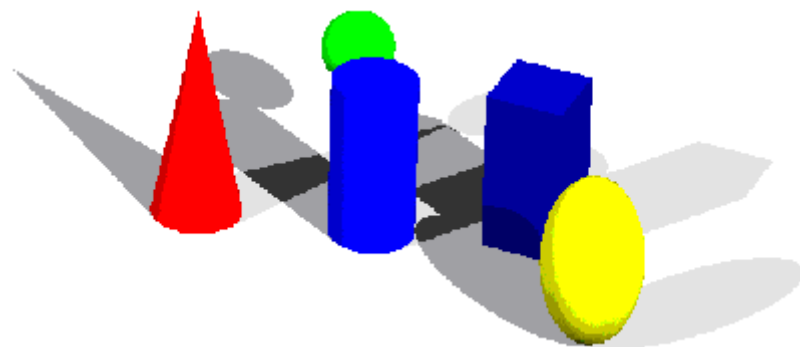
Tel quel, le lancer de rayons ne peut pas restituer un effet de "caustique" tel que l'obtention d'une concentration ou d'une atténuation énergétique en certaines zones lors de la propagation de la lumière à l'intérieur d'une loupe. En effet, cet algorithme de rendu ne modélise que très imparfaitement les transferts d'énergie (la notion d'énergie par unité de surface n'est pas modélisée correctement).

On ne peut pas restituer la diffusion de la lumière dans les matériaux (en particulier l'air).

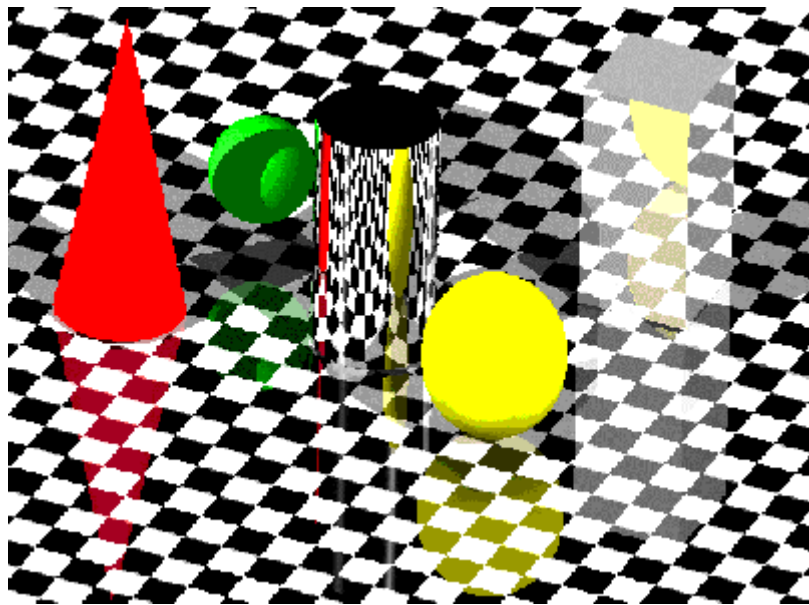
-> Certains éclairages peuvent être non réalistes.

-> Certaines ombres peuvent être non réalistes.

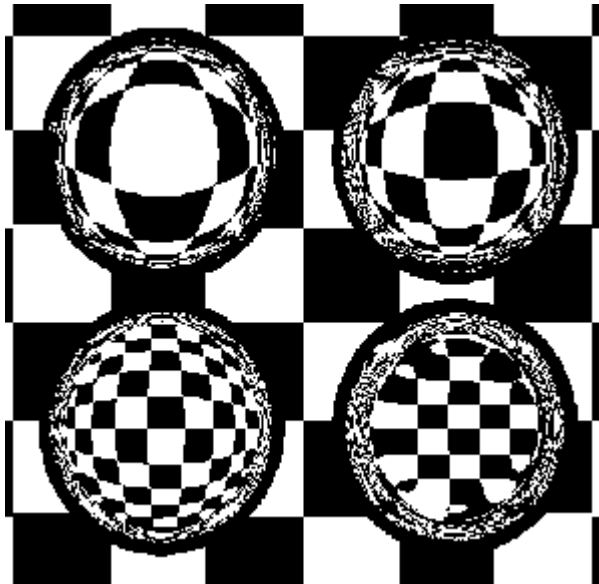
Exemples



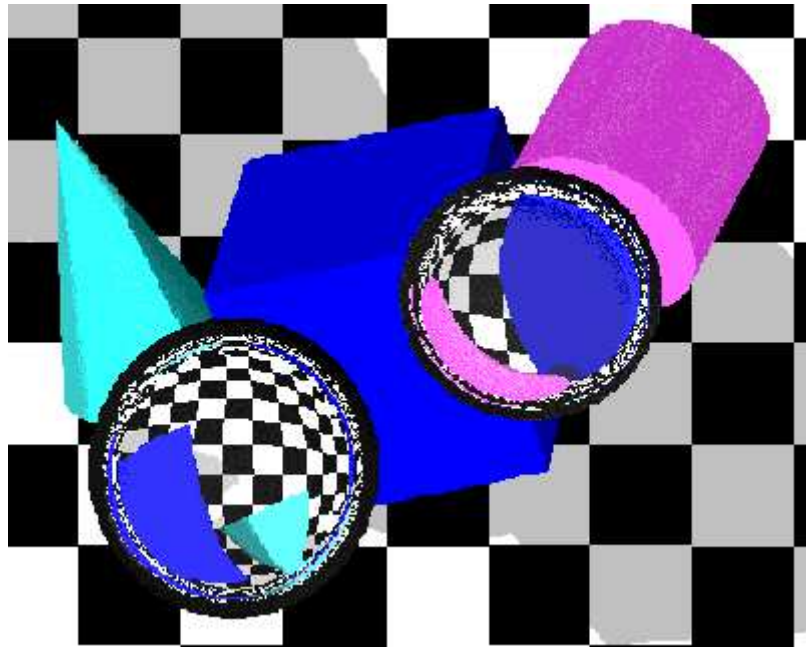
Des ombres portées



Des réflexions



Des transmissions à travers des billes d'indices de réfraction
1.333, 1.5, 2.5 et 4
(de gauche à droite et de bas en haut).



Noter l'inversion des objets.

Quelques problèmes mathématiques

Calcul du rayon réfléchi

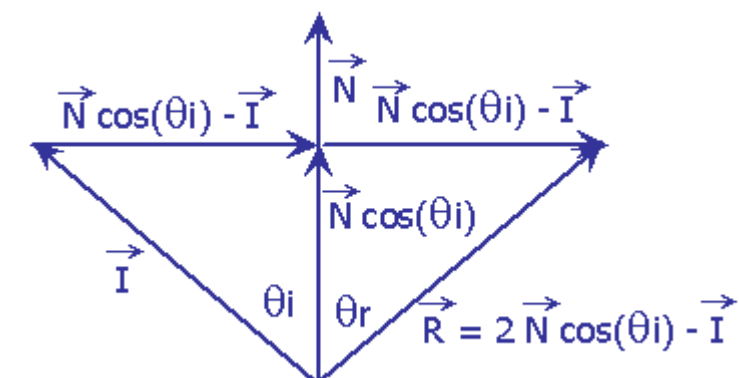
Soient:

- \vec{I} le vecteur normé inverse à la direction d'incidence de la source lumineuse,
- \vec{R} le vecteur normé réfléchi,
- \vec{N} la normale à l'interface orientée dans le matériau où se propage le rayon incident.

Sachant que $\text{abs}(\theta_i) = \text{abs}(\theta_r)$, on établit:

$$\vec{R} = 2 \vec{N} \cos(\theta_i) - \vec{I}$$

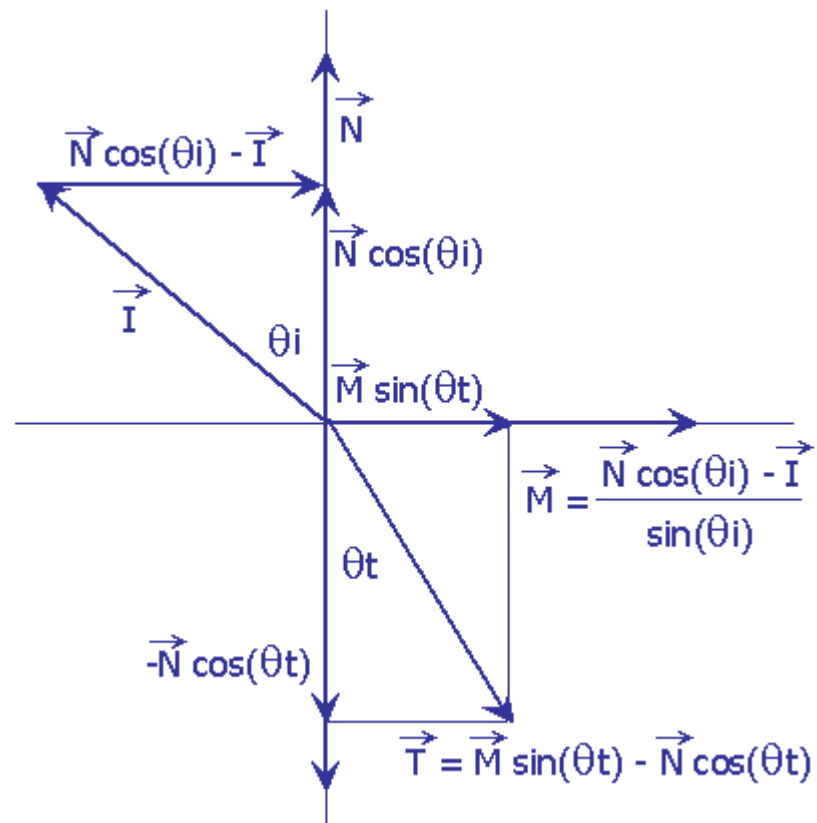
$$\vec{R} = 2 \vec{N} (\vec{N} \cdot \vec{I}) - \vec{I}$$



Calcul du rayon transmis

Soient:

- \vec{I} le vecteur normé inverse à la direction d'incidence de la source lumineuse,
- \vec{T} le vecteur normé transmis,
- \vec{N} la normale à l'interface orientée dans le matériau où se propage le rayon incident,
- \vec{M} le vecteur normé perpendiculaire à \vec{N} situé dans le plan déterminé par \vec{I} et \vec{N} ,
- n_i et n_t les indices de réfraction des milieux incident et de transmission.



$$\vec{T} = \vec{M} \sin(\theta_t) - \vec{N} \cos(\theta_t)$$

$$\vec{M} = \frac{\cos(\theta_i) \vec{N} - \vec{I}}{\sin(\theta_i)}$$

$$\vec{T} = \frac{\sin(\theta_t)}{\sin(\theta_i)} (\cos(\theta_i) \vec{N} - \vec{I}) - \vec{N} \cos(\theta_t)$$

Si on pose $n = \frac{n_i}{n_t}$, on obtient:

$$\vec{T} = (n \cos(\theta_i) - \cos(\theta_t)) \vec{N} - n \vec{I}$$

$$\cos(\theta_i) = \vec{N} \cdot \vec{I}$$

$$\cos(\theta_t) = \sqrt{1 - \sin^2(\theta_t)} = \sqrt{1 - n^2 \sin^2(\theta_i)}$$

$$\cos(\theta_t) = \sqrt{1 - n^2(1 - (\vec{N} \cdot \vec{I})^2)}$$

Ce qui nous conduit pour finir à

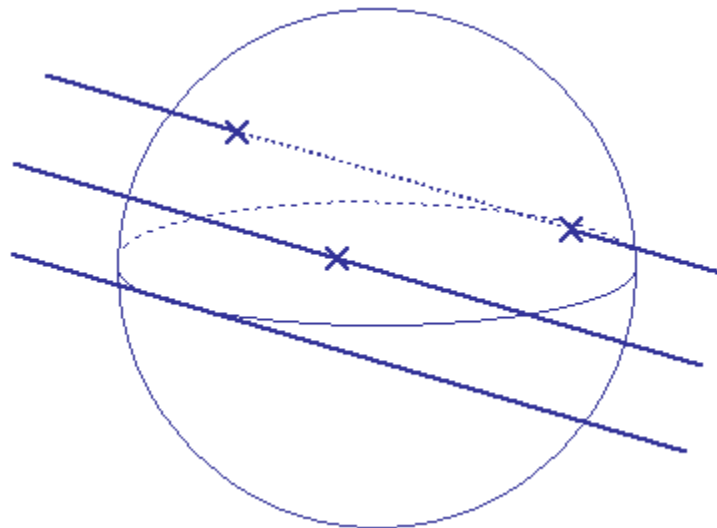
$$\vec{T} = (n(\vec{N} \cdot \vec{I}) - \sqrt{1 - n^2(1 - (\vec{N} \cdot \vec{I})^2)}) \vec{N} - n \vec{I}$$

qui ne contient que des quantités connues.

Calculs d'intersection

Intersection entre une sphère et un rayon

On désire connaître la position de(s) l'intersection(s) entre un rayon lumineux quelconque et une sphère de rayon r centrée sur l'origine.



Le système d'équation à résoudre est formé de l'équation paramétrique de la droite définie par le rayon lumineux et l'équation cartésienne de la sphère:

$$\begin{cases} x = a_1 t + b_1 \\ y = a_2 t + b_2 \\ z = a_3 t + b_3 \\ x^2 + y^2 + z^2 = r^2 \end{cases}$$

Par substitution on obtient:

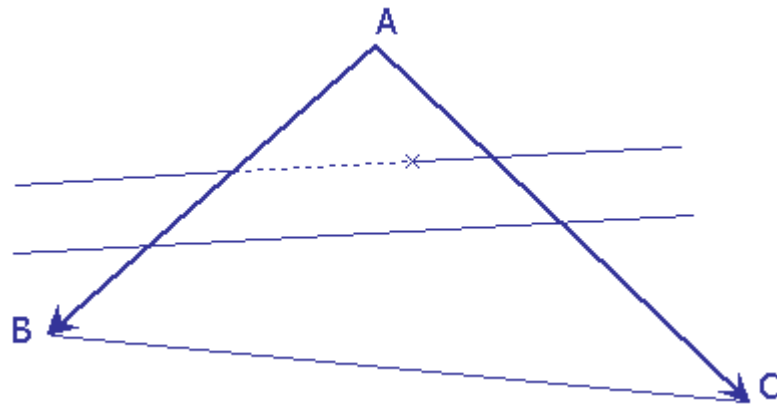
$$(a_1 t + b_1)^2 + (a_2 t + b_2)^2 + (a_3 t + b_3)^2 = r^2$$

$$(a_1^2 + a_2^2 + a_3^2)t^2 + 2(a_1 b_1 + a_2 b_2 + a_3 b_3)t + b_1^2 + b_2^2 + b_3^2 = r^2$$

équation du second degré en t impliquant 0, 1 ou 2 racines et donc 0, 1 ou 2 intersections en fonction des conditions initiales.

Intersection entre une facette triangulaire et un rayon

On désire connaître la position de l'intersection (si elle existe) entre un rayon et une facette triangulaire connue par les coordonnées dans l'espace de ses trois sommets A , B et C .



On a donc: $A = \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$, $B = \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}$ et $C = \begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix}$.

$$B-A = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} x_B - x_A \\ y_B - y_A \\ z_B - z_A \end{pmatrix}, \quad C-A = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_C - x_A \\ y_C - y_A \\ z_C - z_A \end{pmatrix}$$

Le système à résoudre est
$$\begin{cases} x = x_1 u + x_2 v + x_A \\ y = y_1 u + y_2 v + y_A \\ z = z_1 u + z_2 v + z_A \\ x = x_3 t + x_p \\ y = y_3 t + y_p \\ z = z_3 t + z_p \end{cases}$$
 pour déterminer

la position de l'intersection entre le rayon et le plan de la facette.

$$\rightarrow \begin{cases} x_P - x_A = x_1 u + x_2 v - x_3 t \\ y_P - y_A = y_1 u + y_2 v - y_3 t \\ z_P - z_A = z_1 u + z_2 v - z_3 t \end{cases}$$

On doit résoudre ce système de trois équations linéaires à trois inconnues

La solution est:

$$D = -z_1 y_3 x_2 + y_1 z_3 x_2 + x_3 z_1 y_2 + y_3 x_1 z_2 - z_3 x_1 y_2 - x_3 y_1 z_2$$

D doit être différent de 0. Si D est égal à 0, le rayon est parallèle à la facette.

$$u = \frac{-(-y_3 x_2 z_A - y_3 x_P z_2 + y_3 x_2 z_P + y_3 x_A z_2 + x_3 y_2 z_A - x_3 y_2 z_P + x_2 z_3 y_A - x_2 z_3 y_P + x_P z_3 y_2 + x_3 y_P z_2 - x_3 y_A z_2 - x_A z_3 y_2)}{D}$$

$$v = \frac{(-x_3 z_1 y_A - x_3 y_1 z_P + x_3 z_1 y_P + x_3 y_1 z_A + z_3 x_1 y_A + y_1 z_3 x_P + z_1 y_3 x_A - z_3 x_1 y_P - y_3 x_1 z_A - y_1 z_3 x_A + y_3 x_1 z_P - z_1 y_3 x_P)}{D}$$

$$t = \frac{(-x_P z_1 y_2 + x_P y_1 z_2 + x_A z_1 y_2 - x_A y_1 z_2 - x_1 y_2 z_A + x_1 y_2 z_P - x_1 y_P z_2 + x_1 y_A z_2 - x_2 z_1 y_A - x_2 y_1 z_P + x_2 z_1 y_P + x_2 y_1 z_A)}{D}$$

Pour que le point d'intersection soit à l'intérieur du triangle ABC, il faut que:

- u soit positif,
- v soit positif,
- u+v soit inférieur ou égal à 1.

Si on pose $dx = x_P - x_A$, $dy = y_P - y_A$, $dz = z_P - z_A$

$$D = z_1 y_3 x_2 - y_1 z_3 x_2 - x_3 z_1 y_2 - y_3 x_1 z_2 + z_3 x_1 y_2 + x_3 y_1 z_2$$

$$v = \frac{-x_3 y_1 d_z - z_3 x_1 d_y + y_1 z_3 d_x + x_3 z_1 d_y + y_3 x_1 d_z - z_1 y_3 d_x}{D}$$

$$u = \frac{-y_3 d_x z_2 + y_3 x_2 d_z - x_2 z_3 d_y + d_z z_3 y_2 - x_3 y_2 d_z + x_3 d_y z_2}{D}$$

$$t = \frac{-d_{x_1 y_2} + d_{x_1 z_2} + x_1 y_2 d_z - x_1 d_y z_2 - x_2 y_1 d_z + x_2 z_1 d_y}{D}$$

Optimisations

Les optimisations classiques mises en œuvre autour de l'algorithme du lancer de rayons portent sur trois facteurs clés:

- Le calcul des intersections -> Amélioration des tests d'intersection
- Le nombre de calculs d'intersection -> Diminution de ce nombre
- Le nombre de rayons tracés -> Diminution du nombre de rayons primaires et diminution du nombre rayons secondaires

Le calcul des intersections

Le test d'intersection rayon-objet est l'opération de base de tout algorithme de lancer de rayon. Son optimisation pour les différents objets pris en compte permet des gains de temps substantiels.

Représentation mémoire des objets

Tout objet sera représenté sous la forme d'un objet canonique (exemple: le cube de centre O et de rayon 1.0 pour tous les parallélépipèdes rectangles) auquel on applique une matrice de transformation M en coordonnées homogènes.

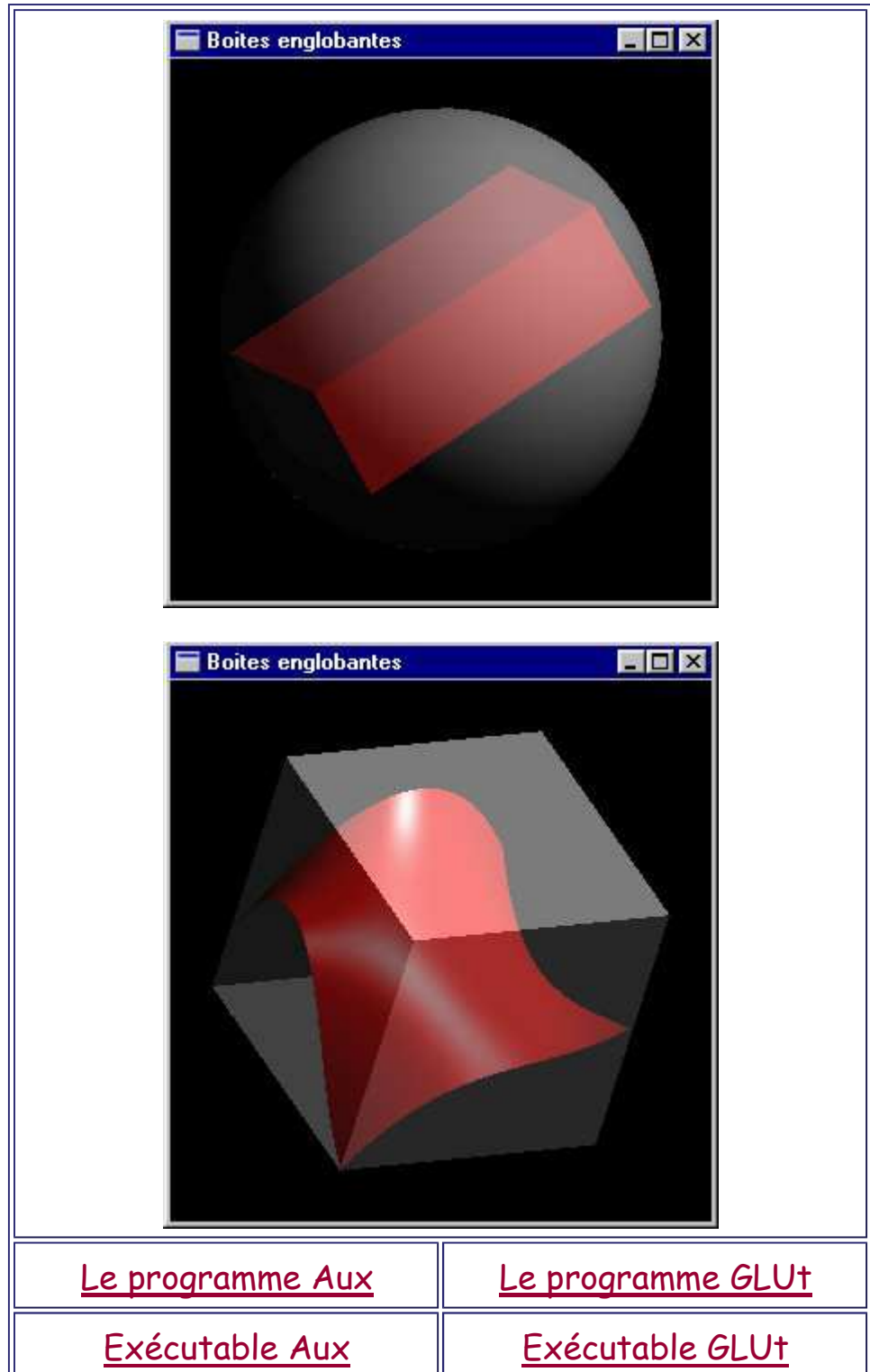
Test d'intersection avec un rayon R:

- Appliquer à R la matrice M^{-1} pour obtenir le rayon R' ,
- Déterminer l'intersection entre R' et l'objet canonique,
- S'il y a une intersection, calculer sa position et les rayons réfléchis et transmis
- Appliquer la matrice M pour déterminer ces rayons dans l'espace de modélisation.

Boîtes englobantes

La réalisation d'un test d'intersection pour un objet canonique peut entraîner une charge de calcul importante (par exemple, pour le tore il n'y a pas de solution analytique mais uniquement des solutions numériques par approximations successives).

En englobant strictement chaque objet dans une "boite" géométrique simple (test d'intersection simple), définie en coordonnées globales, on pourra réaliser une économie substantielle car un rayon qui ne touche pas la boite ne pourra pas toucher l'objet qu'elle contient.

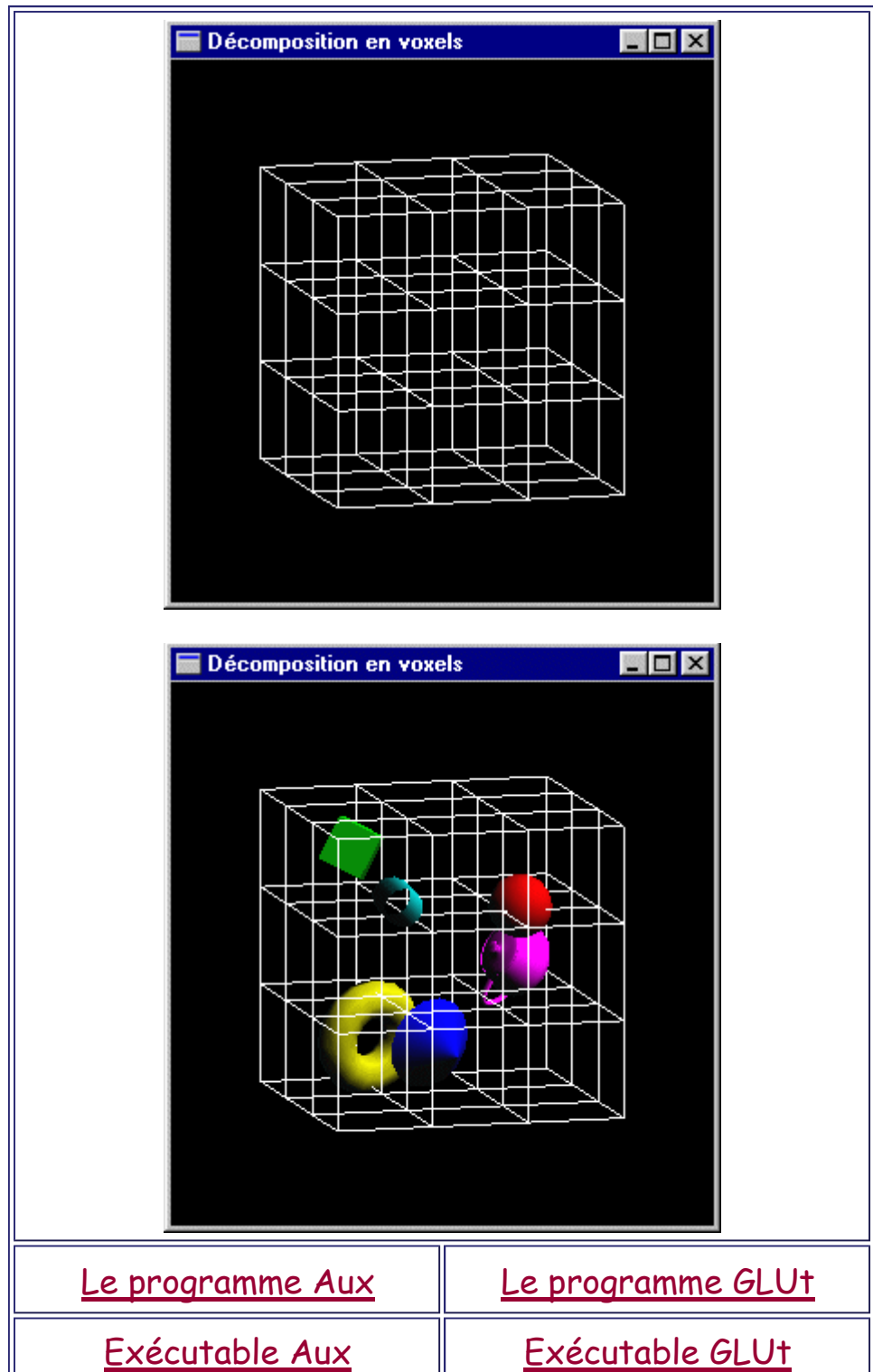


Le nombre de calculs d'intersections

L'acquisition de connaissances sur la position des objets à l'intérieur de la scène peut éviter la réalisation obligatoire d'une recherche systématique d'intersection lors du test d'un rayon vis à vis d'une scène.

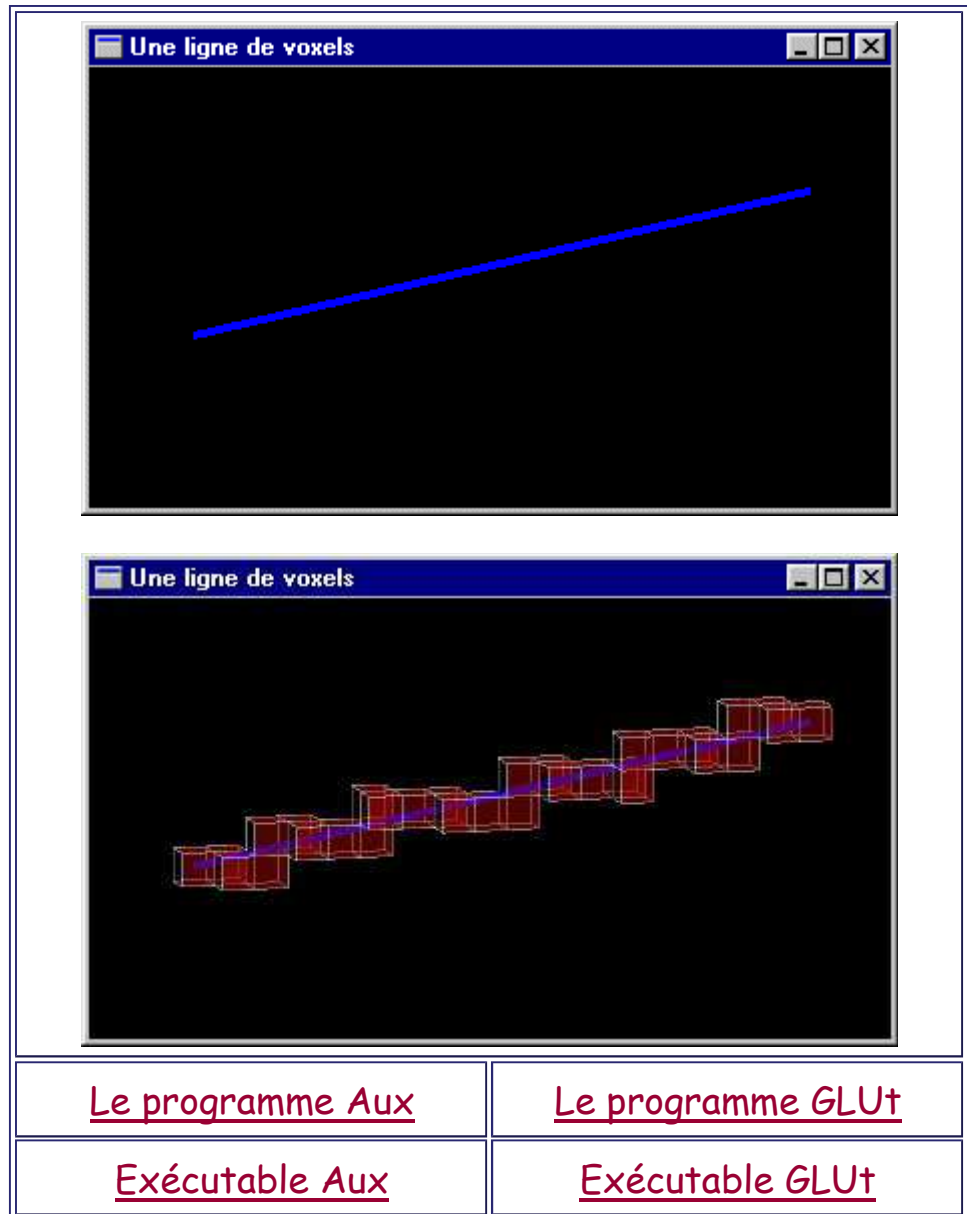
Décomposition en voxels

Une technique classique consiste à décomposer l'espace en voxels et à associer à chaque voxel v la liste des objets partiellement ou totalement inclus dans v .



Plutôt que de parcourir séquentiellement tous les objets de la scène à la recherche de l'intersection la plus proche de la source d'un rayon, on analyse les listes d'objets des voxels, voxel après voxel sur le trajet du rayon depuis le voxel d'émission.

Sitôt qu'une intersection est trouvée, il n'est plus nécessaire de parcourir d'autres voxels car toute autre intersection est forcément plus éloignée du point d'émission.



Le nombre de rayons tracés

Une solution simple pour diminuer le nombre de rayons primaires consiste à effectuer préalablement au lancer de rayon une projection des objets de la scène sur l'écran de visualisation. On saura ainsi quels sont les pixels en lesquels aucun objet n'apparaît.

Le problème de la diminution du nombre de rayons secondaires est complexe et ne sera pas traité ici.

Conclusion

Le lancer de rayons est correct du point de vue géométrique pour le placement des rayons lumineux à l'intérieur d'une scène.

Il est en revanche beaucoup moins correct du point de vue énergétique car les phénomènes que l'on désire rendre relèvent de transports d'énergie, aspect que le lancer de rayons ne traduit que partiellement. Ces phénomènes sont beaucoup plus complexes que ceux que l'on peut modéliser par de simples calculs géométriques.

Un avantage déterminant du lancer de rayons pour l'obtention d'images photoréalistes est toutefois que cet algorithme permet de rendre les réflexions et les transparences.

Par rapport à un algorithme du Z-Buffer + Gouraud ou Phong, les images sont de bien meilleure qualité, mais les temps de calcul sont généralement considérablement plus longs (multipliés par 100, 1000 ou plus).