| Ray Tracing |
| :---: |
| Kadi Bouatouch |
| IRISA |
| Email: kadi@irisa.fr |
| $=$ IRISA |

## Introduction to Computer Graphics

## Context

- Representation and visualization of things which do not exist
- Create 3D virtual worlds
- Choose the model of the objects we want to represent
- Visualize them
- Choose a representation model
- Animate these worlds


## Interest in 3D virtual worlds

- Computer aided design
- Manufacturing
- Movies
- Entertainement: video games
- Data Visualization
- Virtual Reality
- Working without risk
- Simulation




## Rendering Engine

## - Real-time

- Z-buffer
- 10 images per second >> 0,1s
- 25 images per second >> 0,04 s
- 60 images per second >> 0,016 s
- 120 images per second >> 0,008 s
- Non real-time
- Ray Tracing, Path Tracing, Photon Mapping
- Radiosity


## Diffuse Reflection

- Lambert Reflection (Diffuse):
- Part of the incident light penetrates the object then arises from the object with the same intensity in all the directions
$-I_{d}=K_{d} \cdot I_{\text {source }} \cdot \cos (N, L) / d^{2}$
- $\mathrm{K}_{\mathrm{d}}$ : diffuse color of the object
- $I_{\text {source }}$ : intensity of the light source
- $N$ : normale at point $P$
- L : light direction
- $d$ : distance between the light source and point $P$
- Related to the microscopic roughness of the object
- The more $\mathrm{K}_{\mathrm{d}}$ is high, the more important is diffusion


## Illumination Model

- Expresses the light intensity at a point due to:
- The light sources giving rise to :
- diffuse and specular reflections
- shadows
- The reflectance and the transmittance of the object
- $I_{r}=I_{d}+I_{S}$
$-I_{d}$ : diffuse intensity
- $I_{\mathrm{s}}$ : specular intensity


## Diffuse Reflection

- $I_{d}=K_{d} . I_{\text {source }} \cdot \cos (N, L) / d^{2}$
- $d=$ distance of the light source to the lit point $P$



## Specular Reflection

- Phong Model:
- reflection by the object's surface of the part of the incident light that did not penetrates into the object; it depends on the view direction V
$-I_{s}=K_{s} \cdot I_{\text {source }} \cdot \cos ^{n}(R, V) / d^{2}$
$\mathrm{K}_{\mathrm{s}}$ : specular color of the object
- Isource : intensity of the light source
- R : direction of ideal specular reflection
- V : view direction
- n : shininess or roughness
- n high: shiny, tight reflection cone
-n small: mat object, wide reflection cone
- For glass, $\mathrm{n}=200$
- Gives the size of highlight


## Specular Reflection

- $I_{s}=K_{s} \cdot I_{\text {source }} \cdot \cos ^{n}(R, V) / d^{2}$

Light source
Viewer

—IRISA

The Phong Illumination Model

$$
I_{\text {local }}=\sum_{i=0}^{\text {noLLum }} I_{i} \times \frac{v i s(i)}{d_{i}{ }^{2}} \times\left(k_{d}\left(\vec{N} \cdot \overrightarrow{L_{i}}\right)+k_{s}\left(\overrightarrow{R_{i}} \cdot \vec{V}\right)^{n}\right)
$$

$K_{s}$ : specular color ( $R, G, B$ )
$K_{d}$ : diffuse color (R,G,B)
$\mathrm{L}_{\mathrm{i}}$ : lighting direction of source i
$I_{i}$ : intensity of source $i(R, G, B)$


The Phong Reflection Model


- IRISA


## Some Remarks ...

- If multiple sources, sum their contributions
- Several directions have to be known
- The normal to the objects, the light source directions and the view direction:
- the directions $L, N$ et $R$ are coplanar
- The angle ( $\mathrm{L}, \mathrm{N}$ ) is equal to ( $\mathrm{N}, \mathrm{R}$ )
- $\mathrm{K}_{\mathrm{s}}$ and $\mathrm{K}_{\mathrm{d}}$ are 3-component vectors:
- Red, Green, Blue
- $I_{d}+I_{s}=\left(k_{d} \cdot \cos (N, L)+k_{s} \cdot \cos ^{n}(R, V)\right) \cdot I_{\text {source }} / d^{2}$
- The term between () = BRDF


## Some Remarks ...

- $I_{\mathrm{s}}=\mathrm{ks} .<\mathrm{N}, \mathrm{H}>\mathrm{n} . \mathrm{I}_{\text {source }}$
- If the surface is perfectly specular, $n$ is very large
- $<\mathrm{N}, \mathrm{H}>\mathrm{n}$ is not negligible only for $(\mathrm{N}, \mathrm{H})=0$
- Thus $\mathrm{Ir}=\mathrm{ks}$. $\mathrm{I}_{\text {source }}$
- $(\mathrm{N}, \mathrm{H})=0$ means that the incident and reflection angles are equal
- Only 1 reflected ray: because we assume the surface perfectly specular


## To sum up

- Division of the reflectance into 2 components
- Diffuse
- one coefficient $\mathrm{k}_{\mathrm{d}}$
- Specular
- one coefficient $\mathrm{k}_{\mathrm{s}}$
- one coefficient n (shininess)

Approximation de la réflectance

| •Paramètres: |
| :---: |
| $-K_{d}=0.25$ |
| $-K_{s}=0.75$ |
| $-n=50.0$ |

IRIs A

Approximation de la réflectance


- IRISA


Approximation de la réflectance


Recap: Different Light Transports


## Ambient Term

-The indirect diffuse component due to multiple reflections is supposed to be the result of the diffuse reflection of an ambient term la
$-I_{i d}=k_{d} \cdot I_{a}$


- la is the same for all the surfaces


## Principle


$\qquad$
Normal
Normal
Reflected ray
Shadow ray
Refracted ray

- IRISA


## Principle

- Trace a primary ray passing through a pixel
- P : intersection point
- Compute the contribution of the sources to $P$ by tracing shadow rays toward the light sources
- If a shadow ray intersects an opaque object between $P$ and the light source then $P$ is shadowed
Compute the contribution to P of other points within the scene by tracing secondary rays: reflected and refracted
A reflected ray is traced only if the material is specular
- A refracted ray is traced only if the material is transparent
- A secondary ray intersects the scene at a point P'
- Again compute the contribution of the sources to $P^{\prime}$ by tracing
shadow rays toward the light sources
- Repeat the process
- Each ray brings its contribution to the luminance of a point



## Principle

2D Example
right $=$ towards $x$ up
$\Theta=$ frustum half-angle
$d=$ distance to view plane
$P 1=P 0+d^{\star}$ towards $-d^{\star} \tan (\Theta)^{*}$ right $P 2=P 0+d^{\star}$ towards $+d^{\star} \tan (\Theta)^{*}$ right
$\mathrm{P}=\mathrm{P} 1+(\mathrm{i} / \text { width }+0.5)^{*}(\mathrm{P} 2-\mathrm{P} 1)$
$=\mathrm{P} 1+(\mathrm{i} / \text { width }+0.5)^{*} 2^{*} \mathrm{~d}^{*} \tan$
$(\Theta){ }^{*}$ right


Ray: $P=P 0+t V$
$V=(P-P 0) /\|P-P 0\|$

## Ray Generation

- Pinhole camera
for ( $\mathrm{x}=0$; $\mathrm{x}<\mathrm{xres}$; $\mathrm{x}++$ )
for ( $y=0 ; y<y r e s ; y++$ )
\{
$\underline{\mathbf{d}}=\underline{\mathbf{f}}+2(\mathrm{x} / \mathrm{xres}-0.5) \cdot \underline{x}$
$+2(y / y r e s-0.5) \cdot \underline{\underline{y}}$
$\frac{d}{r}=\frac{d}{d} /\left|\frac{d}{d}\right| ; / /$ Normalize
color= ray_cast ( $r$, scene,depth) ; write_pixel(x,y,color);
\}

-路


## REFLECTION



- IRISA


## REFRACTION


$n_{1} \sin (\alpha)=n_{2} \sin (\beta)$

- IRISA


## Refraction: Using Snell's Law

$$
\frac{\sin \alpha}{\sin \beta}=\frac{\eta_{2}}{\eta_{1}}=\eta_{21}
$$

- Using this law it is possible to show that:
$T=-\eta_{12} E+N\left(\eta_{12} \cdot \cos \alpha-\sqrt{1+\eta_{12}^{2} \cdot\left(\cos ^{2} \alpha-1\right)}\right)$
- Note that if the root is negative then total internal reflection has occurred and you just reflect the vector as normal


## Ray-Tracing: Pseudocode

- For each ray $\mathbf{r}$ from eye to pixel, color the pixel with the value returned by ray_cast(r , scene,depth):
ray_cast(r, scene,depth)
\{
If(depth $>$ Max_Depth) $\{$ color $\leftarrow$ black $\}$ else \{

If (intersection(r,scene)) \{
$\mathbf{p} \leftarrow$ point_of_intersection $(\mathbf{r}$, scene); $\mathbf{u} \leftarrow \operatorname{reflect}(\mathbf{r}, \mathbf{p})$;
$\mathbf{v} \leftarrow \operatorname{refract}(\mathbf{r}, \mathbf{p}) ;$
color $\leftarrow$ phong_direct $(\mathbf{p}, \mathbf{r})+$
$\mathrm{k}_{\mathrm{s}} \times$ ray_cast(u, scene, depth +1 ) +
kt $\times$ ray_cast( $\mathbf{v}$, scene, depth +1 ); $\}$ else color $\leftarrow$ background_color ;
\}
return(color);
\}

## Pseudocode Explained

- $\mathbf{p} \leftarrow$ point_of_intersection( $\mathbf{r}$, scene);
- Compute $\mathbf{p}$, the point of intersection of ray $\mathbf{r}$ with the scene
- $\mathbf{u} \leftarrow \operatorname{reflect}(\mathbf{r}, \mathbf{p}) ; \mathbf{v} \leftarrow \operatorname{refract}(\mathbf{r}, \mathbf{p}) ;$
- Compute the reflected ray $\mathbf{u}$ and the refracted ray v using Snell's Laws


## Pseudocode Explained

- phong(p,r)
- Evaluate the Phong reflection model for the ray $\mathbf{r}$ at point $\mathbf{p}$ on surface s , taking shadowing into account
- $\mathrm{k}_{\mathrm{s}} \times$ ray_cast(u,scene,depth)
- Multiply the contribution from the reflected ray u by the specular color $\mathrm{k}_{\mathrm{s}}$ for surface s containing p. Only (specular-to-specular)* light transport is handled. Ideal specular (mirror) reflection
- $\mathrm{k}_{\mathrm{t}} \times$ ray_cast( $\mathbf{v}$,scene,depth)
- Multiply the contribution from the refracted ray $\mathbf{v}$ by the specular-refraction coefficient $k_{t}$ for surface $s$.

Only (specular-refraction)* light transport is handled

- IRISA


## About Those Calls to ray_cast()...

- The function ray_cast() calls itself recursively
- There is a potential for infinite recursion
- Consider a "hall of mirrors"
- Solution: limit the depth of recursion
- A typical limit is five calls deep
- Note that the deeper the recursion, the less the ray's contribution to the image, so limiting the depth of recursion does not affect the final image much


## About Those Calls to ray_cast()...

- Another solution
- Ei: direct lighting at point Pi
- Ks: vector (R,G,B)
- Kt: scalar ranging between 0 and 1
- Contribution of the red path
-I: Intensity due to this ray path
$=K s 0 \cdot(K t 1(K s 4 \cdot E 7+E 4)+E 1)$
$=\mathrm{Ks} 0 . \mathrm{Kt1} . \mathrm{Ks} 4 . \mathrm{E} 7+\mathrm{Ks} 0 . \mathrm{Kt1} . \mathrm{E} 4+\mathrm{Ks} 0 . \mathrm{E} 1$
- Stop tracing rays when the cumulative product Ks.Kt... is below a certain threshold

Example

H1 : bisecting line of angle S P3 P2 H2 : bisecting line of angle S P2 P1
H 1 : bisecting line of angle S P1 O
Idai : intensity due to direct lighting dai : intensity due to direct lighting and the ambient term for point Pi
Idai $=$ kdi . la
kdi . Is . $\cos (\mathrm{Li}, \mathrm{Ni})$
+ksi . Is $\cdot \cos (\mathrm{Ni}, \mathrm{Hi})^{\mathrm{n}}$
$13=$ Ida3
$12=$ Ida2 + ks2 .13
$\mathrm{I} 1=\mathrm{Ida} 1+\mathrm{ks} 1 . \mathrm{I} 2$


## Reflections

- If only one reflected ray is considered, then raytracing will only handle perfect mirrors.

- IRISA


## Reflections

- Glossy reflections (multiple reflected rays) blur the reflection.



## Glossy Reflections

- We need to integrate the color over the reflected cone.
- Weighted by the reflection coefficient in that direction.



## Translucency

- Likewise, for blurred refractions, we need to integrate around the refracted angle.

-IRISA



## Calculating the integrals

- How do we calculate these integrals?
- Two-dimensional of the angles and ray-depth of the cone.
- Unknown function -> the rendered scene.
- Use Monte-Carlo integration


## Shadows

- Ray tracing casts shadow from a point light source.
- Many light sources are illuminated over a finite area.
- The shadows between these are substantially different.
- Area light sources cast soft shadows
- Penumbra
- Umbra



## Soft Shadows

- Umbra - No part of the light source is visible.
- Penumbra - Part of the light source is occluded and part is visible (to a varying degree).
- Which part? How much? What is the Light Intensity reaching the surface?
- IRISA


## Pros and Cons of Ray Tracing

- Advantages of ray tracing
- All the advantages of the Phong model
- Also handles shadows, reflection, and refraction
- Disadvantages of ray tracing
- Computational expense
- No diffuse inter-reflection between surfaces
- Not physically accurate
- Other techniques exist to handle these shortcomings, at even greater expense! =IRISA


## An Aside on Antialiasing

- Our simple ray tracer produces images with noticeable "jaggies"
- Jaggies and other unwanted artifacts can be eliminated by antialiasing:
- Cast multiple rays through each image pixel
- Color the pixel the average ray contribution
- An easy solution, but it increases the number of rays, and hence computation time, by an order of magnitude or more


## Intersection

## Intersection

## Sphere

## Principle

- The scene is supposed to be expressed in the world coordinate system (WCS).
- It may be: A set of independent objects
- Purpose: intersect a scene with a ray whose equation is given by :
- $P=P_{0}+t . D$
- where:

Po is the ray origin;
$\mathrm{D}=(\mathrm{dx}, \mathrm{dy}, \mathrm{dz})$ is the direction vector of the ray ;
$t>0$

- Intersection result $=\{\mathrm{ti} / \mathrm{ti}$ is a value of t corresponding to an intersection point $\}$
- Only the closest point to the ray origin is used to compute shading and secondary shot rays
$d_{0}$ : Orthogonal distance between the ray and the $P=P$ the sphere of radius $r$ and center $C$
- $P=\left(X \quad Y Z_{0}\right) D=(d x, d y d z)$
- $P_{0}=\left(X_{0}, Y_{0} Z_{0}\right) D=(d x, d y, d z)$
- If $d_{0}{ }^{2}<=r^{2}$, then the ray intersects the sphere

Intersection points $=$ solutions of
$\left\|P_{0}-C\right\|^{2}+2 t \cdot\left(P_{0}-C\right) \cdot D+t^{2} \cdot\|D\|^{2}=r^{2}$
$\mathrm{d}_{0}$ is evaluated by minimizing the
distance $d$ between $C$ and a point $P$ on the ray.
This gives:
$d^{2}=\left\|P_{0}+t \cdot D-C\right\|^{2}=\left\|P_{0}-C\right\|^{2}+$
$2 t .\left(P_{0}-C\right) \cdot D+t^{2}\|D\|^{2}$
2t. ( $\left.P_{0}-C\right) . D+t^{2} \cdot\|D\|^{2}$

- By setting to 0 the derivative of $\mathrm{d}^{2}$, we obtain :
$t=\left(\left(P_{0}-C\right) . D /\|D\|^{2}\right)=-\left(P_{0}-C\right) . D$
After substitution: $\mathrm{d}_{0}{ }^{2}=\left\|\mathrm{P}_{0}-\mathrm{C}\right\|^{2}-$
$((P-C) . D)^{2}$



## Intersection

## Axis-aligned Parallelepiped

- Faces: perpendicular to the axes of the world coordinate system.
- First, the intersections between the ray and the faces $x=x 1$ and $x=$ x2 are computed.
Two values of $t$ are then obtained
- $\mathrm{t} 1=(\mathrm{x} 1-\mathrm{x} 0) / \mathrm{dx}$ and $\mathrm{t} 2=(\mathrm{x} 2-\mathrm{x} 0) / \mathrm{dx}$.
- Interval: [ $\mathrm{Ix}, \mathrm{Mx}$ ] = [ $\min (\mathrm{t} 1, \mathrm{t} 2)$ ), $\max (\mathrm{t} 1, \mathrm{t} 2)$ ]
- Same processing applied to the faces perpendicular to the $y$ and $z$ axes. Two other intervals: [ ly, My ] and [ $\mathrm{Iz}, \mathrm{Mz}$ ]
- The result is then an intersection interval given by
[I, M ] = [ max( lx, ly, Iz ), min( Mx, My, Mz ) ]
- If I <= M then the ray intersects the parallelepiped bounding volume, otherwise it does not intersect it
- Closest intersection point: $\mathrm{t}=1$


## Intersection

## Polyhedron

- Polyhedron = set of pairs of
 parallel faces
- Ni: normal to a pair of faces
- A pair of parallel faces is
 called slab

- IRISA


## Intersection

## Polyhedron

- The value of $t$ corresponding to the intersection between the ray and this face is computed by substituting the ray equation into that of the plane :
$t=-(d+N . P O) / N . D$
$t=\alpha_{i} * d+\beta i$
- For each slab i, N=Ni and $\quad===\rightarrow \quad \alpha i=\frac{-1}{N i \bullet D}$
- Given a slab i, these values are the
$\beta i=\frac{-N i \bullet P 0}{N i \bullet D}$ same for all the object bounding volumes


## Intersection

## Cylinder

- The cylinder : intersection between an infinite height cylinder and the subspace delimited by two planes which equations are $z=0$ and $z=h$
The intersection between the ray an the infinite height cylinder is first performed. This yields a first interval [t1,t2]
- The intersection with the two planes gives a second interval [ $\mathrm{t} 3, \mathrm{t} 4$ ].
- The final intersection interval $[I, M$ ] results from the combination of these two intervals (as for the parallelepiped).



## Intersection

## Cylinder: continued

- obtaining [ $\mathrm{t} 1, \mathrm{t} 2$ ]
- The equation of the infinite height cylinder :
$-x^{2}+y^{2}=r^{2}$
- Substituting the ray equation in this equation we obtain:
$t^{2} \cdot\left(d x^{2}+d y^{2}\right)+2 t \cdot\left(x_{0} \cdot d x+y_{0} \cdot d y\right)+\left(x_{0}{ }^{2}+y_{0}{ }^{2}-r 2\right)=0$
- Solving this equation gives the interval [ $\mathrm{t} 1, \mathrm{t} 2$ ].
- obtaining [t3, t4]
_ Let $A$ and $B$ the two values of $t$ resulting from the intersection with the two planes
$A=-z_{0} / d z$ and $B=\left(h-z_{0}\right) / d z$
- We get :
$\mathrm{t} 3=\min (A, B)$ and $t 4=\max (A, B)$
— IRISA


## Intersection

## Cone

- Intersection: performed in the LCS of the cone
- Cone: intersection between an infinite height cone and the subspace delimited by two planes, the equations of which are $z=0$ and $z=h$
- Intersection between the ray and the infinite height cone is first performed.
- The equation of this cone is given by :
h2 . $(x 2+y 2)-r 2 . z 2=0$



## Intersection

## Cone

- Substituting the ray equation in this equation yields an interval [ t1, t2 ].
- Then the planes are in their turn intersected to give a second interval [ $t 3, t 4$ ] such that :

$$
\mathrm{t} 3=\min (\mathrm{A}, \mathrm{~B}) \text { and } \mathrm{t} 4=\max (\mathrm{A}, \mathrm{~B})
$$

- where $A=-z 0 / d z$ and $B=(h-z 0) / d z$.
- The final interval is the combination of these two intervals (as for the cylinder)


## Intersection

## Polygon

- Several ray-polygon intersection methods have been proposed in the literature.
- Only two of them are presented.
- For all these methods, the intersection process consists of two steps :
- First step: Ray-Plane intersection test
- the goal of the first step is to perform the intersection between the ray and the plane containing the polygon
- Second step: Inside - Outside test
- the second step tests if the resulting point is inside or outside the polygon.
- IRISA


## Intersection - Triangle

## - Barycentric coordinates

- Non-degenerate triangle ABC
$\underline{P}=\lambda_{1} \underline{A}+\lambda_{2} \underline{B}+\lambda_{3} \underline{C}$
$-\lambda_{1}+\lambda_{2}+\lambda_{3}=1$
$-\lambda_{3}=\operatorname{area}(\mathrm{APB}) / \operatorname{area}(\mathrm{ACB})$,
$-\lambda_{2}=\operatorname{area}(\mathrm{APC}) / \operatorname{area}(\mathrm{ACB})$,

$-\lambda_{3}=\operatorname{area}(\mathrm{CPB}) /$ area(ACB),
- Area(APB)= $\left.\operatorname{det}(P \vec{A}, P \vec{B})=\frac{1}{2}|P \vec{A} \times P \vec{B}|=\frac{1}{2}|P \vec{A}| P \vec{B} \right\rvert\, \sin (\hat{P})$
- $\lambda_{i}$ greater or equal than zero


## Intersection

## - Polygon: Snyder's method

- Ray-triangle intersection: extension to a polygon.
- Let Pi be the vertices of a triangle and Ni the associated normals which are used for normal interpolation across the triangle.
- Normal to the triangle: $\mathrm{N}=(\mathrm{P} 1-\mathrm{P} 0) \times(\mathrm{P} 2-\mathrm{P} 0)$
- A point $P$ lying on the triangle plane satisfies :

$$
P . N+d=0 \quad \text { where } d=-P 0 . N \text {. }
$$

- To intersect a ray $P=O+t$. $D$ with a triangle, first compute the t parameter of the intersection between the ray and the triangle plane
$t=(d-N . O) / N . D$.


## Intersection

Polygon: Snyder's method

- Projecting the triangle into any other plane, except one that is orthogenal to the triangles plane will not change the barycentric coordinates of the triangle.
- This allows to simplify computations, since we can choose any of the coordinate system's three axis-
aligned planes to project our triangle, thus throwing away one of the three coordinates and reducing the barycentric equations to $\mathrm{R}^{2}$.
- For reasons of numerical stability we want to choose the dominant axis of the triangles normal for the projection.
- An index i0 is computed: equal either to 0 if $|N x|$ is maximum (i.e. the $x$ axis is dominant) or to 1 if $|\mathrm{Ny}|$ is maximum or to 2 if $|\mathrm{Nz}|$ is maximum.


## Intersection

## Intersection

## Polygon: Snyder's method

Let i1 and i2 (i1, i2 $\in\{0,1,2\}$ ) be two unequal indices different from i0 mpute the in and 12 components of the intersection point
$\mathrm{li} 1=\mathrm{Oi} 1+\mathrm{t} . \mathrm{Di} 1$ and $\mathrm{li} 2=\mathrm{Oi} 2+\mathrm{t} . \mathrm{Di} 2$

- The inside-outside test can be performed by computing scalars $B 0, B 1$ and $\mathrm{Bi}=[(\mathrm{Pi}+2$ $\mathrm{Bi}=[(\mathrm{Pi}+2-\mathrm{Pi}+1) \times(\mathrm{I}-\mathrm{Pi}+1)] \mathrm{iO} /[\mathrm{N}] \mathrm{iO}$
- The Bi are the barycentric coordinates of the point where the ray intersects the triangle plane.
I is inside the triangle if and only if $0 \leq \beta \leq 1$ for $i\{0,1,2\}$
The interpolated normal at point $I$ is given by

$$
\mathrm{N}^{\prime}=\beta 0 . N 0+B 1 . \mathrm{N} 1+\beta 2 . \mathrm{N} 2 .
$$

. Snyder's method can be easily extended up to polygons.

- The main idea is to consider a polygon as a union of triangles


## Marchal's method

I is the ray-plane intersection point
The Pi are transformed to the two dimensional coordinates system ( $u, v$ ) whose origin is vertex
The
The plane of this coordinates system is the polygon plane.
The inside-outside test determines if an edge The inside-outside test determines if an edge occur when the $u$ components of Pi and $\mathrm{Pi}+1$ have different signs ).
If So, and if POI < POM then I is inside the polygon,
 else it is outside.
On the other hand, if none of the edges intersec the $v$ axis, then I lies outside the polygon

## Intersection

## Marchal's method

- The interpolated normal at point I is given by : $\mathrm{NI}=(\mathrm{POI} / \mathrm{POM}) \cdot \mathrm{NM}+(1-\mathrm{POI} / \mathrm{POM}) \cdot \mathrm{NO}$
- where the normal NM at point M is given by :
$\mathrm{NM}=(\mathrm{PiM} / \mathrm{PiPi}+1) \cdot \mathrm{Ni}+1+(1-\mathrm{PiM} / \mathrm{PiPi}+1) . \mathrm{N}$

- and $\mathrm{Ni}, \mathrm{Ni}+1$ are the normals at point Pi and $\mathrm{Pi}+1$. $\mathrm{PiPi}+1$ is the intersected edge.


## Bounding box

- To reduce the amount of ray-object intersections, its is absolutely necessary to use a hierarchical data structure .
- This data structure is a tree of bounding volumes.
- Bounding volumes are simple geometric objects which fit around the objects.
- They are chosen to be simple to intersect with a ray, such as spheres or parallelepipeds that have faces perpendicular to the axes.



## Bounding Volume

## Different kinds of bounding Volume

- Parallelepiped
- For the sake of speed up, the faces of this bounding volume are perpendicular to the axes of the World Coordinates System.
- Its perspective projection onto the screen plane is often used to filter the primary rays (rays starting at the eye location).
- Sphere and Ellipsoid
- They may be used to filter the reflected and refracted rays and those directed to the light sources.
- Polyhedron
- Intersection of slabs: a slab is a pair of parallel faces


## Bounding Volume Hierarchy

- Organize objects into a tree
- Group objects in the tree
- based on spatial relationships
- Each node in the tree
contains a bounding box of all
the objects below it


## Bounding Volume Hierarchy (BVH)

- Determining optimal BVH structure is NP-hard problem
- Heuristic approaches:
- Cost models (minimize volume or surface area)
- Spatial models
- Categories of approaches:
- Top down
- Bottom up



## Median Cut BVH Construction

## Top down approach:

- Sort objects by position on axis
- cycle through $x, y, z$
- use center of bounding box
- Insert tree node with half of objects on left and half on
 right


## Median Cut BVH Construction

1. $L=\{l i s t ~ o f ~ b o u n d i n g ~ v o l u m e ~ n u m b e r s ~\} ~$
2. Choose widest slab:
dmax[2]-dmin[2] or dmax[1]-dmin[1] dmax[2] (In this example : max width = dmax[1]-dmin[1]
3. Then choose slab of max width
4. Sort the bounding volumes wrp to increasing
dmin[number_of_widest_slab]
5. We get a sorted list $\bar{L}=\{1, \overline{5}, 3,2,4\}$
6. Split L into two sub-lists L 1 and L 2
7. We get : $\mathrm{L} 1=\{1,5,3\} \quad \mathrm{L} 2=\{2,4\}$

8. Go to 1 with $\mathrm{L}=\mathrm{L} 1$ then $\mathrm{L}=\mathrm{L} 2$

Leaf $=$ one or more objects

## Bottom up BVH Construction

- Add objects one at a time to tree
- Insert to subtree that would cause smallest increase to area


## Bottom up BVH Construction

- Add objects one at a time to tree

- Insert to subtree that would cause smallest increase to area



## Bottom up BVH Construction

- Add objects one at a time to tree
- Insert to subtree that would cause smallest increase to area



## Bottom up BVH Construction

- Add objects one at a time to tree
- Insert to subtree that would cause smallest increase to area
- IRISA


## Intersection Test Using the BVH

- Once the hierarchy of bounding volumes has been built, the ray-scene intersection test is performed as follows.
- The hierarchy is searched from the root to the leaves.
- During this search, at a node N , the associated bounding volume is checked for an intersection with the current ray.
- If the bounding volume of N is intersected, those of its children node are in their turn checked for an
intersection.
- This process is repeated recursively and ends up at the leaf nodes.
- Else, if the bounding volume of N is not intersected by the ray, the associated subtree is left out, that is, it is not searched.
$\qquad$


## Spatial Subdivision

- The rectangular bounding volume of the scene is subdivided into 3 D cells
- Each cell contains a few objects of the scene
- When a ray enters a cell, we check the objects within this cell for an intersection with the ray
- If the intersection process ends up with success then no need to check the rest of the objects
- If the ray fails to hit any object in the cell then it moves to the next 3D cell
- Repeat the process


## Spatial Subdivision

- Two procedures
- A procedure which performs a spatial subdivision of the scene into 3D cells, each of them containing a small portion of the database
- A second procedure which determines the next cell along a ray


## Spatial Subdivision

- Two procedures


Uniform Grid





## Spatial Subdivision

- Different kinds of subdivision



## Uniform Grid

## Uniform Grid

## Ray Traversal Algorithm: Classical Method

- Let $G[i][j][k]$ be the $3 D$ array representing the 3D grid

Let $P$ the point where the ray leaves the current cell and $D$ the ray
direction
P is the outgoing point

- Let $w$ be the axis perpendicular to the face which contains $P$
- Let $u(x, y$ or $z)$ be the index ( $i, j$ or $k$ ) of the current cell corresponding to

If $\mathrm{Dw}>0$ then the index $u$ of the next cell is $u=u+1$, the other indices are unchanged

- Else it is: $u=u-1$
- Example :
- If $w=z$ then $u=k$
- If $=z=0$ ne

If $\mathrm{Dz}>0$ then the index of the next cell along the ray is $\mathrm{k}=\mathrm{k}+1$, while the
other indices do not change
If
if $D z>0$, or $G[i i l j i] k ~ G[i]] j i j][k] ~ t h e n ~ t h e ~ n e x t ~ c e l l ~ a l o n g ~ t h e ~ r a y ~ i s ~ G[i j] j][k ~+1] ~$
if $\mathrm{Dz}>0$, or $\mathrm{G}[\mathrm{i}[\mathrm{j}][\mathrm{k}-1]$ if $\mathrm{Dz}<0$

## Uniform Grid

## Ray Traversal Algorithm: Classical Method

(2,0

## Uniform Grid

- Ray Traversal Algorithm: Amanatide's Algorithm



## Uniform Grid

Ray Traversal Algorithm: Amanatide's Algorithm
Initialization

- Ray equation: $P=P 0+t$. $D$
- Identify the voxel containing the ray origin $O$
- If $O$ is outside the grid, find the point through which the ray enters the grid and determine the adjacent voxe
- X, Y and Z : voxel indices
- StepX, stepY and stepZ : initialized to 1, incremented o decremented as the ray crosses the voxel boundaries
tx, ty and tz: values of t corresponding to the points resulting from the intersection between the ray and 3 faces of the initial voxel
- tDeltaX, tDeltaY and tDeltaZ : distance travelled by the ray between two successive faces perpendicular to the $x, y$ and $z$ faces respectively


## Uniform Grid



## Uniform Grid

## - Advantages?

- easy to construct
- easy to traverse
- Disadvantages?
- may be only sparsely filled
- geometry may still be clumped (say, densely grouped)


## Non Uniform Grid

- The rectangular bounding volume of the scene is recursively sliced: - either simultaneously by 3 planes perpendicular to the $x, y$ and $z$ axes: Octree
- or by one plane at a time perpendicular to an axis: Kd-tree, Bsp tree
- or by one plane at a time non necessary perpendicular to an axis: Bsp tree
- Each slicing plane divides a space (a 3D cell) into two subspaces (3D cells)
- The subdivision process stops either when a cell contains partially or totally a minimum number of objects or the maximum subdivision level is reached for each axis
- The result is a linear array of rectangular cells or a binary tree or an octree
- Each cell is represented by a data structure containing a pointer to the objets partially or totally within it


## Non Uniform Grid

- Subdivide until each cell contains no more than
- $n$ elements, or maximum depth $d$ is reached

- IRISA


## Non Uniform Grid

- Advantages?
- grid complexity matches geometric density
- Disadvantages?
- more expensive to traverse (especially octree)



## Non Uniform Grid: Kd-Tree

- Subdivide only 1 dimension
- Do not subdivide at the center
- Which axis to pick?
- What point on the axis to pick?
- One heuristic:
- Sort objects on each axis
- Pick point corresponding to "middle" object
- Pick axis that has "best" distribution of objects
$-\mathrm{L}=\mathrm{n} / 2, \mathrm{R}=\mathrm{n} / 2$ (ideal)
- Realistically,
- minimize (L-R) and
- L approx. n/2, R approx. n/2

-IRISA



## kD-Trees: Data Structure

```
Struct KdTreeNode {
    union {
        u_int axis; // Both, x or y or z split plane (0,1,2), 3 for leaf
        float value; I/ Interior, split position
        u_int nPrims; // Leaf
    }
    union{
        Primitive *onePrimitive; // Leaf
        Primitive nPrims; I/ Leaf
    }
    KdTreeNode *LeftChild; // interio
    KdTreeNode *RightChild; II interior
}
# IRISA
```

KD-Tree: Traversal


## KD-Tree: Traversal

## Input: a tree and a ray

-Search for the first intersected primitive in the tree
-Traversal: start from the root
-Use of a stack
-First range of $t$, [tmin,tmax]: associated with the scene bounding box -Internal node encountered: ray is classified wrp to the splitting plane
-If range lies entirely in one side of the plane, traversal moves to the appropriate child
-If the range straddles the plane, traversal will continue to the first child hit by the ray while the second child is pushed onto the stack along with its range [tmin,tmax]
-Traversal proceeds down the tree, occasionally pushing items onto the stack, until a leaf node is reached.

## KD-Tree: Traversal

```
```

kd-search( tree, ray )

```
```

kd-search( tree, ray )
(global-tmin, global-tmax) = intersect( tree.bounds, ray )
(global-tmin, global-tmax) = intersect( tree.bounds, ray )
{
{
search-node( tree.root, ray, global-tmin, global-tmax )
search-node( tree.root, ray, global-tmin, global-tmax )
}
}
search-node( node, ray, tmin, tmax )
search-node( node, ray, tmin, tmax )
{
{
if( node.is-leaf )
if( node.is-leaf )
search-leaf( node, ray, tmin, tmax )
search-leaf( node, ray, tmin, tmax )
else
else
search-split( node, ray, tmin, tmax )
search-split( node, ray, tmin, tmax )
}

```
```

}

```
```


## KD-Tree: Traversal

```
search-split( split, ray, tmin, tmax ) {
    a = split.axis
    thit = ( split.value - ray.origin[a] )/ ray.direction[a]
    (first, second) = order( ray.direction[a], split.left,split.right )
    if( thit >= tmax or thit < 0 )
    search-node( first, ray, tmin, tmax )
    else if( thit <= tmin)
        search-node( second, ray, tmin, tmax )
        else{
            stack.push( second, thit, tmax )
            search-node( first, ray, tmin, thit )
        }
}

\section*{KD-Tree: Traversal}
```

search-leaf( leaf, ray, tmin, tmax )
// search for a hit in this leaf
if( found-hit and hit.t < tmax )
succeed( hit)
else
continue-search(leaf, ray, tmin, tmax )
}
continue-search(leaf, ray, tmin, tmax ){
if( stack.is-empty )
fail()
else{
(n, tmin, tmax) = stack.pop()
search-node( n, ray, tmin, tmax )
}
}

```
-Remark
If stack empty, then no intersection along the ray and the search terminates


\section*{Kd-tree traversal: Observation}
- Eliminate stack operations
- How?
- If the traversal reaches a leaf and fails to find a hit:
- Restart the search at the root
- With tmin advanced to the end of the leaf
- The first leaf intersected by the modified range is the next leaf that needs to be traversed

\section*{Kd-tree traversal: Restart}
```

continue-search( leaf, ray, tmin, tmax )
{
if( tmax == global-tmax )
fail()
else {
tmin = tmax
tmax = global-tmax
search-node( tree.root, ray, tmin, tmax )
}
}
-IRISA

```


\section*{Kd-tree: Backtrack}
- In the traditional, a node pushed onto the stack is always the other (second) child of one of the current node's ancestors
- Thus, possible to reach the parent of the node atop the stack by following a chain of parent links (which we can store in the nodes of the tree) from the current node.
- If we again employ the tactic of advancing tmin to the end of the last leaf visited, then we will be able to recognize the appropriate parent as the closest ancestor that has a nonempty intersection with the remaining (tmin; tmax) range.
- Increase per-node storage

\section*{KD-Backtrack}
```

continue-search(leaf, ray, tmin, tmax ){
if( tmax == global-tmax )
fail()
else {
tmin = tmax
tmax = global-tmax
backtrack( leaf.parent, ray, tmin, tmax )
} {
}
backtrack( split, ray, tmin, tmax ) {
(t0,t1) = intersect( split.bounds, ray, tmin, tmax )
if( no-intersection)
backtrack( split.parent, ray, tmin, tmax )
else
search(split, ray, t0, t1)
}

```
- Generalization of kd-
trees
- Splitting plane is not
axis aligned
- Used in games:

DOOM


\section*{BSP tree}
- A Binary Space Partitioning (BSP) tree data structure
- Recursive, Hierarchical subdivision of \(n\) dimensional space into convex subspaces.
- BSP tree construction
- Partition a subspace by a hyper-plane that touches the edge of the subspace.
- The result is two new subspaces that can be further partitioned by recursive application of the method.
- A "hyperplane" in an n-dimensional space is an \(n-1\) dimensional object which can be used to divide the space into two half-spaces.
- IRISA

\section*{BSP tree}
- example:
- In three dimensional space,
the "hyperplane" is a plane
- In two dimensional space,
it is a line.
- BSP trees are extremely versatile, because they are powerful sorting and classification structures
- Hidden surface removal
- Ray tracing hierarchies
- Solid modeling
- Robot motion planning.


\section*{Building a BSP tree}
- Given a set of polygons in three dimensional space, we would like to build a BSP tree which contains all of the polygons.
- The algorithm to build a BSP tree:
- Select a partition plane.
- Partition the set of polygons with the plane.
- Recurse with each of the two new sets.
- The choice of partition plane depends on how the tree will be used, and what sort of efficiency criteria you have for the construction.
- IRISA

\section*{Building a BSP tree}
- For some purposes, it is appropriate to choose the partition plane from the input set of polygons
- Scan-conversion
- Other applications may benefit more from axis aligned orthogonal partitions
- Ray tracing
- Space subdivision.
- It is desirable to have a balanced tree, where each leaf contains roughly the same number of polygons.
- It is desirable to minimize polygon splitting.
- Finding the optimal split is hard, we use a heuristic
- Testing the plane against a small random number of (5-6) polygons for split.

\section*{BSP tree: Partitioning}
- Classify each member of the set with respect to the plane.
- If a polygon lies entirely on one side of the hyper-plane
- It is added to the partition set for the proper side.
- If a polygon spans the plane - keep in the node
- If the polygon intersect the hyper-plane
- Split it as needed and add the parts the proper sets.

\section*{BSP tree: When to stop?}
- The decision to terminate the tree construction is a matter of the specific application.
- Some applications will benefit from termination when the number of polygons in a leaf node is below a maximum value.
- Other methods continue until every polygon is placed in an internal node.
- Another criteria that can be used is the maximum tree depth.

\section*{BSP tree: example}
- One of the most important properties of BSP trees is that it is view independent.
- For example, consider the following case:


\section*{BSP tree: example}
- Splitting the plane using the ordered lines from the input, we get the following:

(1)


BSP tree: example




\section*{BSP tree: example}

\section*{BSP tree: example}
- Now, we can choose several view points, and choose the painting order according to the tree we


\section*{BSP tree: Hidden Surface} Removal
- Probably the most common application of BSP trees is hidden surface removal in three dimension.
- BSP trees provide an elegant, efficient method for sorting polygons via a depth first tree walk. This fact can be exploited in a back to front "painter's algorithm".
- The idea behind the painter's algorithm is to draw polygons far away from the eye first, followed by drawing those that are close to the eye.
- Hidden surfaces will be written over in the image as the surfaces that obscure them are drawn.
- Can assist in 3D clipping.
- Can support Back Face Culling.
- IRISA

\section*{BSP tree: Painting}
- One reason that BSP trees are so elegant for the painter's algorithm is that the splitting of complex polygons is an automatic part of tree construction.
- When building a BSP tree specifically for hidden surface removal, the partition planes are usually chosen from the input polygon set.
- However, any arbitrary plane can be used if there are no intersecting or concave polygons.

\section*{BSP tree: Drawing the scene}
- To draw the contents of the tree:
- Perform a back to front tree traversal.
- Begin at the root node and classify the eye position with respect to the partition plane.
- Draw the subtree at the far child from the eye
- Draw the polygons in this node
- Draw the near subtree.
- Repeat this procedure recursively for each subtree.
- Front to back rendering is also possible.

\section*{BSP tree: Hidden Surface Removal}
- The painting order from V 1 : \(-3,5,1,4 b, 2,6,4 a\)
- The painting order from V 2 : \(-3,5,1,4 b, 2,4 a, 6\)

\section*{BSP tree: Ray Tracing}

\section*{-Accelerating Ray Tracing}
-Rectangular bounding volume of the scene: recursively subdivided
-Subdivision: Splitting planes are axis aligned
-Each splitting plane splits a cell into two equally sized sub-cells
-Choose \(\mathrm{x}, \mathrm{y}\) and z axis one at a time

\section*{Octree}
- Useful for reducing the number of ray-object intersections.
- The bounded 3D world to be ray traced is subdivided into cells of varying size. Each cell contains a list of objects (of approximately the same length) which intersect it.
- Given a ray to be traced, a list of cells intersected by the ray is determined. Intersection calculations are performed only with these objects.
- Furthermore, if the cells may be accessed in the order of advance of the ray, the procedure may terminate once the first intersection is discovered.

\section*{Octree}
- Each node of the tree has eight children, corresponding to halving he space along all of the three axii.
- A node is a leaf if the subspace epresents intersects at most a given number of objects.
- The two basic operations needed for ray tracing octrees are
- Locating the leaf cell containing a given 3D point (point location).
- Locating the next cell intersecting a given ray
- The first is a standard octree traversal The second is
raversal. The second is
with a point along the ray just outside the current cell.

\section*{Octree}
- Octrees ignore the directionality of objects.
- Subdivision is always in predefined directions and places.
- Advantage: Simple construction. Point location is easy.

- Disadvantage: Non-optimal subdivision (large trees).

Octree: example


\section*{Octree: traversal}
1. Determine the first intersection point \(F\) between the ray and the scene's axis aligned bounding box (SAABB)
. Push \(\mathbf{F}\) along the normal to the face containing it,
3. Pushing consists in adding to the P's coordinates a value deltax (resp. deltay, deltaz) which is equal to half the length of the \(x\) side (resp. deltay, deltaz) which is equ
(resp. y, z) of the smallest cell.
4. Search for the cell (containing \(\mathbf{F}\) ) in the tree
5. If no intersection in the cell, compute outgoing point \(\mathbf{P}\)
6. Push \(\mathbf{P}\) along the normal to the cell's face containing it
7. The results is another point \(P^{\prime}\)
8. Search for the cell (containing \(\mathbf{P}^{\prime}\) ) in the tree
9. Go to 1 until intersection

Remark: If \(\mathbf{P}\) is on an edge or a vertex of a cell, push it simultaneously in the directions of the normals to the faces sharing it
Octree: traversal

Uniform vs. Adaptive Subdivision
- Uniform: too much traversed empty cells
- Adaptive: less
empty
cells

—IRISA



\section*{Cells \& Portals}

Teller and Sequin's Approach
(1) Decompose space into convex cells: use walls as splitting polygons
(2) For each cell, identify its boundary edges into two sets: opaque or portal
(3) Precompute visibility among cells
(4) During viewing (eg, walkthrough phase), use the precomputed Potentially Visible polygon Set (PVS) of each cell to speed-up rendering
- IRISA

\section*{Cells \& Portals: Space \\ Subdivision}

Input Scene:


Spatial subdivision: Generated by computing a k-d tree of the input faces
-IRISA
Determining Adjacent Information



\section*{Eye-to-Cell Visibility}

The eye-to-cell visibility of any observer is a subset of the cell-to-cell visibility for the cell containing the observer
```

