Raytracing II

Proseminar Einführung in die Computergrafik
Lehrstuhl für Informatik VIII, RWTH Aachen

Elmar Hoffmann

2. September 2001

Inhaltsverzeichnis

1 Reflexion und Brechung

Eine Stärke des Raytracings gegenüber einfacheren Renderverfahren ist die realistische Darstellung von Licht-Reflexionen und -Brechungen.

Im folgenden werden Reflexions- und Brechungsarten von jeweils idealisierten Materialtypen vorgestellt, in der Realität sind Materialien jedoch nicht ideal. Daher wird bei der praktischen Berechnung das einfallende Licht in Anteile entsprechend der Materialeigenschaften des Objekts zerlegt und die jeweiligen idealen Reflexions- und Brechungsarten mit den Anteilen berechnet.

1.1 Reflexion

Reflexion tritt nicht nur in Form von Spiegelungen auf, sie ist Voraussetzung für das Sehen von Objekten. Das von Objekten reflektierte Licht ist es, welches den Farbeindruck des Objekts erzeugt, wenn es in das menschliche Auge gelangt.

1.1.1 Perfekte Reflexion

Die perfekte Reflexion, engl. *Specular Reflection*, findet man bei glatten, glänzenden Materialien. Es findet keinerlei Absorption statt, das Licht wird ausschließlich reflektiert, die Photonen "prallen" von der Oberfläche "ab".

Daraus ergibt sich das Brechungsgesetz[1], welches besagt, daß der, jeweils zum Normalvektor der Oberfläche gemessene, Einfallswinkel des Lichtstrahls gleich dem Ausfallswinkel ist, siehe Abbildung 1.1.

Im folgenden soll die Formel zur Berechnung des reflektierten Lichtstrahls auf Grund des Brechungsgesetzes hergeleitet werden, dabei seien:

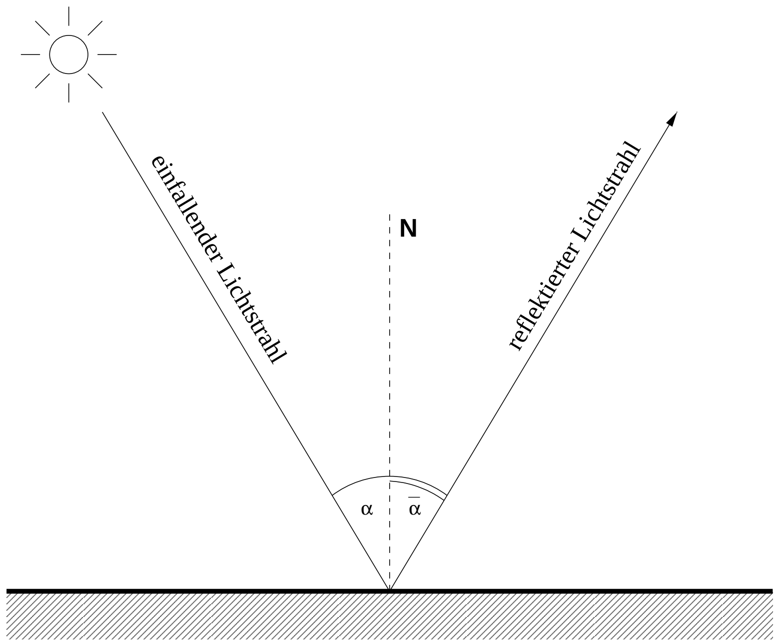
N Normalvektor der Oberfläche
I Richtungseinheitsverktor des einfallenden Lichtstrahls
R Richtungseinheitsverktor des reflektierten Lichtstrahls

Bei gegebenenem **N** und **I** ist nun **R** gesucht:

$$\mathbf{R} = s\mathbf{I} + t\mathbf{N} \tag{1.1}$$

$$\alpha = \bar{\alpha} \tag{1.2}$$

[1] Physikalisch betrachtet lässt sich das Brechungsgesetz aus dem Huygens-Prinzip über Elementarwellen oder aus dem Fermatschen Minimalprinzip herleiten, siehe Gerlach, Grosse [2]



Einfallswinkel = Ausfallswinkel:

$$\alpha = \bar{\alpha}$$

Abbildung 1.1: Reflexion

Aus (1.2) folgt mit $\cos\alpha = -\mathbf{I}\cdot\mathbf{N}$ und $\cos\bar{\alpha} = \mathbf{R}\cdot\mathbf{N}$:

$$\begin{aligned}
\cos\alpha &= \cos\bar{\alpha} && (1.3)\\
-\mathbf{I}\cdot\mathbf{N} &= \mathbf{R}\cdot\mathbf{N} && (1.4)\\
&= \mathbf{N}\cdot(s\mathbf{I} + t\mathbf{N}) && (1.5)\\
&= s(\mathbf{I}\cdot\mathbf{N}) + t(\underbrace{\mathbf{N}\cdot\mathbf{N}}_{=1}) && (1.6)\\
&= s(\mathbf{I}\cdot\mathbf{N}) + t && (1.7)\\
& && (1.8)
\end{aligned}$$

Durch willkürliches Setzen von $s = 1$ erhält man:

$$t = -2(\mathbf{I}\cdot\mathbf{N}) \qquad (1.9)$$

Somit erhält man für den Richtungsvektor des reflektierten Strahls:

$$\mathbf{R} = \mathbf{I} - 2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} \qquad (1.10)$$

1.1.2 Diffuse Reflexion

Die diffuse Reflexion findet man bei rauhen, matten Materialien. Es findet keinerlei Reflexion statt, das Licht wird ausschließlich absorbiert und die dabei aufgenommene

Energie wieder in Form von Photonen abgestrahlt[2].

Da die neu abgestrahlten Photonen in keiner Beziehung zu den absorbierten stehen, geschieht die Abstrahlung in beliebiger (zufälliger) Richtung. Aus diesem Grund wirkt das Objekt matt, es bilden sich keine Glanzpunkte.

Der Einfallswinkel des absorbierten Lichts spielt lediglich für die Menge an absorbierter Energie eine Rolle, vergleichbar mit der hohen Intensität der Sonne Mittags und der geringen Abends.

Daraus ergibt sich, daß sich die Amplitude des abgestrahlten Lichts proportional zum Cosinus des Einfallswinkels θ verhält, siehe Abbildung 1.2.

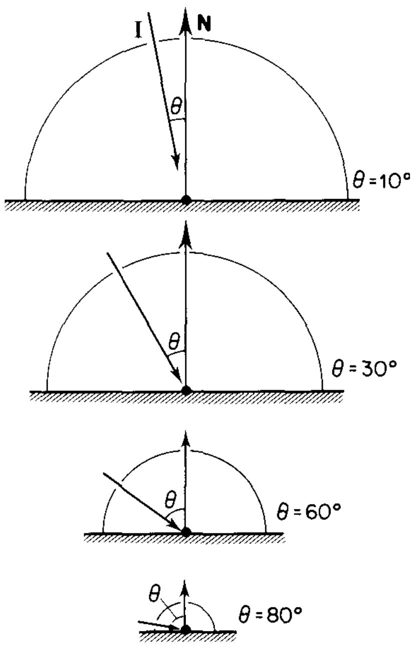


Abbildung 1.2: Diffuse Reflexion

1.2 Brechung

An transparenten Objekten tritt Lichtbrechung auf, analog der Reflexion kann man zwischen ideal perfekter und diffuser Brechung unterscheiden.

1.2.1 Perfekte Brechung

Die perfekte Brechung, engl. *Specular Refraction*, findet man bei klaren, durchsichtigen Materialien.

[2]Der Farbeindruck entsteht hierbei dadurch, daß die Photonen mit einer durch die Materialeigenschaften bestimmten Frequenz abgestrahlt werden.

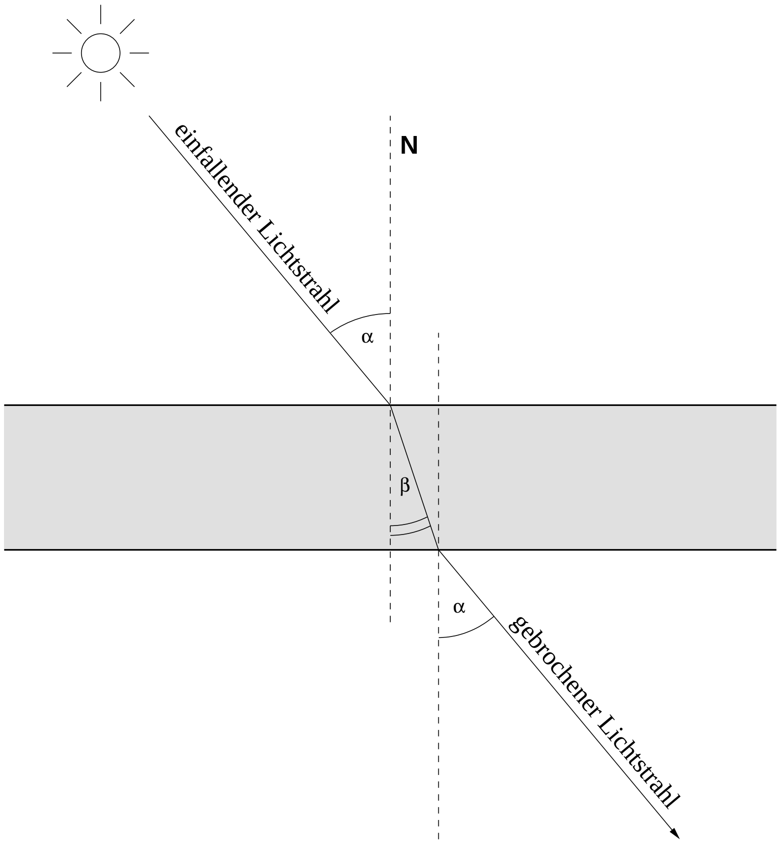
Brechung tritt an der Grenzfläche zwischen zwei unterschiedlich transparenten Materialien auf, die Richtungsänderung des passierenden Lichtstrahls beruht auf der unterschiedlichen Lichtgeschwindigkeit in den Materialien[3].

Man definiert als eine Materialeigenschaft den den Brechungsindex n_x als Verhältnis der Lichtgeschwindigkeit im Material c_x zu der im Vakuum c_0:

$$n_x = \frac{c_x}{c_0}$$

Beispielsweise ist der Brechungsindex von Wasser 1.33, der von Glas liegt je nach Glassorte im Bereich von 1.5 bis 1.7.

Abbildung 1.3 zeigt ein transparentes Objekt und einen beim Ein- und Austritt gebrochenen Lichtstrahl und nebenstehend das Brechungsgesetz.



$$\frac{\sin\alpha}{\sin\beta} = \frac{c_1}{c_2} = \frac{n_2}{n_1}$$

Abbildung 1.3: Brechung

[3]Auch dies lässt sich wiederum aus dem Huygens-Prinzip über Elementarwellen oder aus dem Fermatschen Minimalprinzip herleiten, siehe Gerlach, Grosse [2]

1.2.2 Diffuse Brechung

Die diffuse Brechung findet man bei trüben, durchscheinenden Materialien. Analog zur diffusen Reflexion wird das Licht absorbiert, auf Grund der Transparenz des Objekts jedoch auch in sein Inneres – und somit durch es hindurch – abgestrahlt. Die Amplitude des abgestrahlten Lichts ist wiederum proportional zum Cosinus des Einfallswinkels.

2 Anti-Aliasing und Supersampling

2.1 Aliasing

Aliasing ist ein Effekt der immer dann auftreten kann, wenn kontinuierliche Größen in diskrete umgewandelt werden.

Abbildung 2.1 zeigt zwei Sinuskurven unterschiedlicher Frequenz, die an den durch Pfeile gekennzeichneten Stellen abgetastet werden. Während die durch Kreise gekennzeichneten diskreten Werte von Kurve (a) diese noch hinreichend genau wiedergeben, um die Kurve erkennen zu lassen, lassen die diskretisierten Werte von Kurve (b) eine Sinuskurve niedrigerer Frequenz erkennen. Man sagt, die niedrige Frequenz ist ein *Alias* der hohen, darin liegt der Ursprung des Begriffes Aliasing.

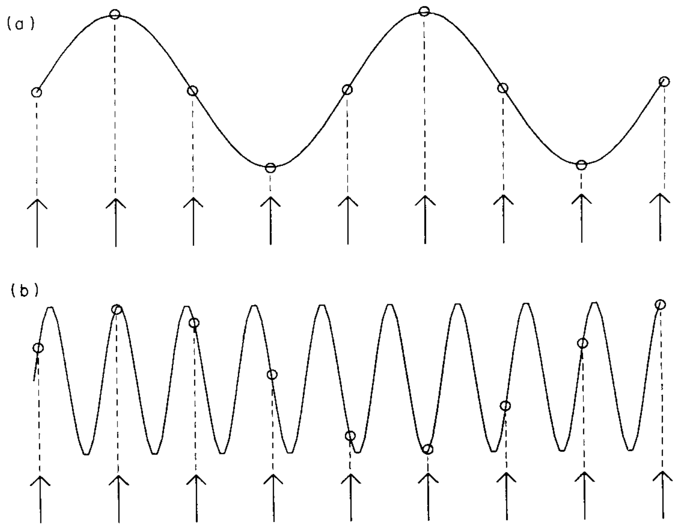


Abbildung 2.1: Aliasing

Wie im Beispiel zu sehen, tritt Aliasing erst ab einer bestimmten Frequenz auf, der sogenannten *Nyquistfrequenz*. Diese leitet sich vom *Nyquist Theorem* ab, welches besagt, daß bei einer Diskretisierung lediglich Signale mit Frequenzen maximal der halben Abtastfrequenz korrekt erfasst werden können.

In der Praxis unterscheidet man zwischen räumlichem und zeitlichen Aliasing.

2.1.1 Räumliches Aliasing

Aus der 2D Grafik bekannt ist der Alias-Effekt der Treppenbildung bei der Diskretisierung einer Linie in das vorgegebene Pixelmuster.

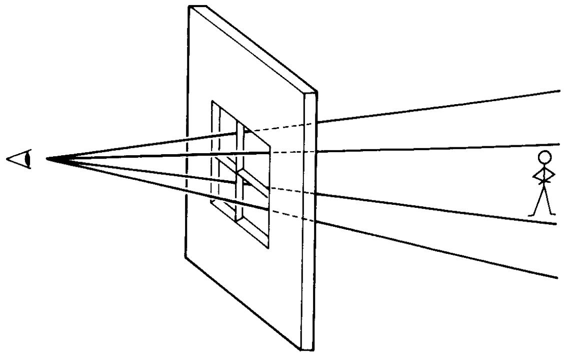
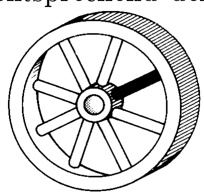


Abbildung 2.2: Räumliches Aliasing

Beim Raytracing tritt dieser Effekt ebenfalls auf, zudem können wie in Abbildung 2.2 gezeigt kleine Objekte oder Details von Texturen unberücksichtigt bleiben, bzw. kleine bewegte Objekte bei Animationen, die in ihrer Bewegung nicht immer von Strahlen getroffen werden plötzlich verschwinden und wieder auftauchen, ein "Popping" genannter, sehr störender Effekt.

2.1.2 Zeitliches Aliasing

Das wohl bekannteste Beispiel für *zeitliches Aliasing* ist ein gefilmtes Rad eines fahrenden Wagens. Die kontinuierliche Drehung des Rades wird dabei entsprechend der Zahl der Bilder pro Sekunde des Films diskretisiert.



Der dabei auftretende Aliasing-Effekt, daß das Rad bei bestimmten Geschwindigkeiten des Wagens zu stehen oder sich rückwärts zu drehen scheint, soll nun anhand des nebenstehend abgebildeten Rades mit einer schwarzen Speiche genauer betrachtet werden.

In Abbildung 2.3 sind drei verschieden schnelle Drehungen dieses Rades mit 6 Abtastungen pro Zeiteinheit diskretisiert worden.

Abbildung 2.3 (a) zeigt eine langsame Drehung, die Rotationsfrequenz liegt unter der *Nyquistfrequenz*, die Rotationsrichtung im Uhrzeigersinn ist eindeutig erkennbar.

Abbildung 2.3 (b) zeigt eine Rotation mit genau der *Nyquistfrequenz*, die Rotationsrichtung lässt sich nichtmehr eindeutig bestimmen.

In Abbildung 2.3 (c) schließlich liegt die Rotationsfrequenz über der *Nyquistfrequenz*, die Rotationsrichtung erscheint fälschlich als gegen den Uhrzeigersinn, auch die Anzahl der Umdrehungen und somit auch die Rotationsgeschwindigkeit erscheinen falsch.

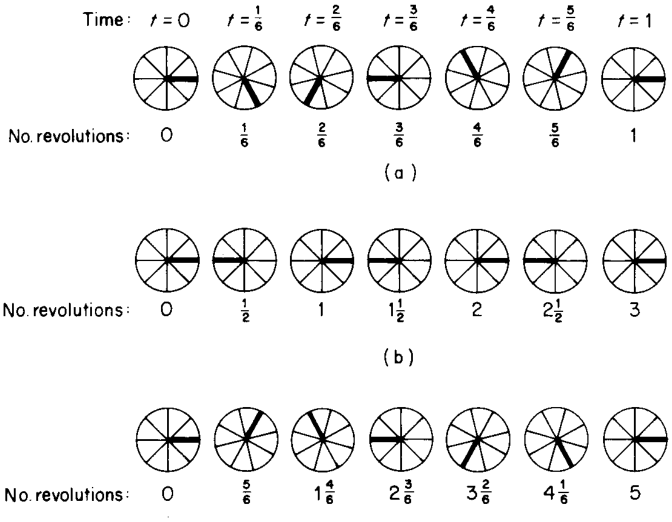


Abbildung 2.3: Zeitliches Aliasing

2.2 Supersampling

Ein einfacher Ansatz, das Aliasing zu bekämpfen, ist das *Supersampling*, hierbei werden zusätzliche Strahlen pro Pixel berechnet. Daraus ergibt sich eine Erhöhung der *Nyquistfrequenz* und somit eine Verringerung der Aliasing-Effekte.

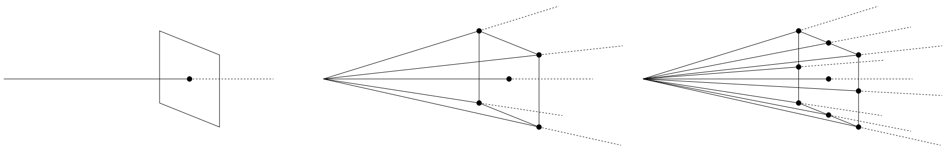


Abbildung 2.4: Verwendung mehrerer Strahlen pro Pixel

Ein großer Nachteil ist die Vervielfachung der Rechenzeit entsprechend der Zahl der Strahlen pro Pixel, desweiteren werden die Aliasing-Effekte zwar verringert, nicht jedoch vermieden.

2.3 Adaptive Sampling

Das *Adaptive Sampling* verwendet analog zum *Supersampling* mehr Strahlen pro Pixel, allerdings lediglich an an den Stellen, wo es nötig erscheint. Dadurch fällt der zusätzliche Rechenaufwand pro Pixel geringer aus als beim *Supersampling*, sprich der größte Nachteil wiegt geringer.

In Abbildung 2.5 ist ein Pixel abgebildet, für den das Adaptive Sampling exemplarisch durchgeführt werden soll. Die Bildebene, auf die "durch" den Pixel geschaut wird, besteht aus 5 verschiedenfarbigen Flächen A bis E.

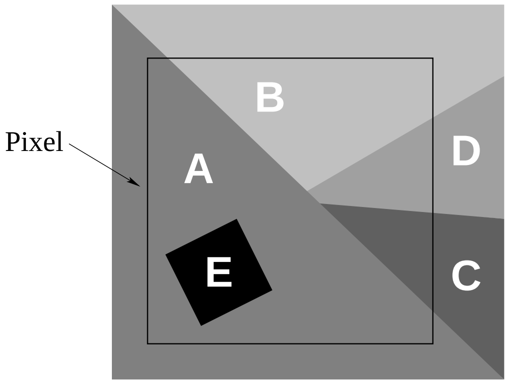
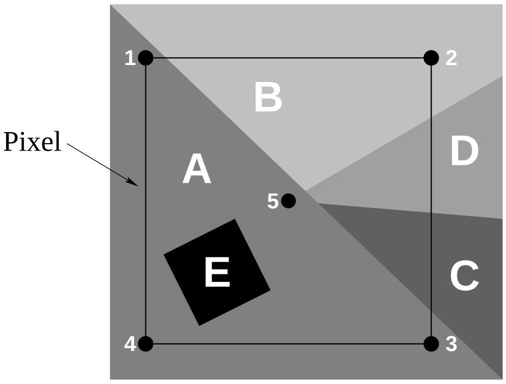


Abbildung 2.5: Ein Pixel vor der Bildebene

Zunächst wird Supersampling mit 5 Strahlen durchgeführt, siehe Abbildung 2.6. Anschließend werden jeweils die ermittelten Farbwerte der Strahlen 1 und 5, 2 und 5, 3 und 5 sowie 4 und 5 miteinander verglichen. Überschreitet die Differenz bei zwei Strahlen eine festgelegte Schranke, wie hier im Beispiel zwischen den Strahlen 2 und 5, so wird im von diesen Strahlen aufgespannten Quadrat erneut Supersampling durchgeführt, siehe Abbildung 2.7. Da die Farbwerte für die Strahlen 2 und 5 bereits vorliegen, müssen nur die Farbwerte der Strahlen 6, 7 und 8 berechnet werden.



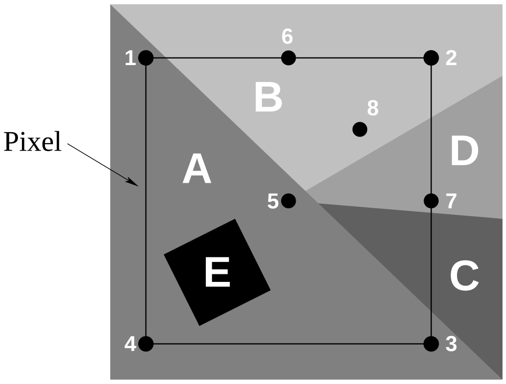
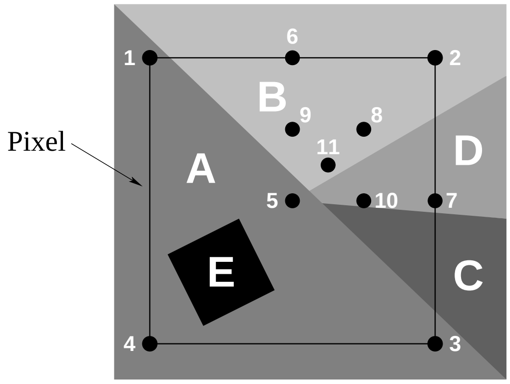


Abbildung 2.6: Supersampling mit 5 Strahlen

Abbildung 2.7: Adaptive Sampling zwischen Strahl 2 und 5

Im nächsten Schritt werden wiederum jeweils die ermittelten Farbwerte der Strahlen 5 und 8, 6 und 8, 2 und 8 sowie 7 und 8 miteinander verglichen. In diesem Fall überschreiten die Differenzen der Farbwerte zwischen den Strahlen 5 und 8 sowie 7 und 8 die festgelegte Schranke und somit wird erneut Supersampling betrieben, siehe Abbildungen 2.8 und 2.9.

Dies wird analog fortgesetzt bis sämtliche Differenzen der Farbwerte die Schranke



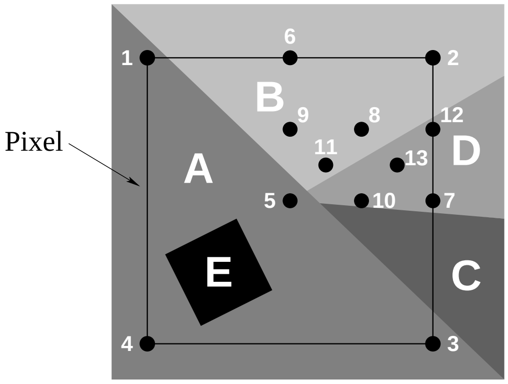


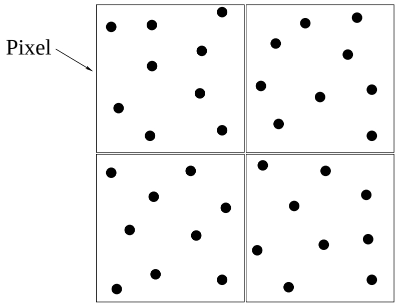
Abbildung 2.8: Adaptive Sampling zwischen Strahl 5 und 8

Abbildung 2.9: Adaptive Sampling zwischen Strahl 7 und 8

unterschreiten oder eine festgelegte maximale Sampletiefe erreicht wird. Letzteres soll im Beispiel an dieser Stelle der Fall sein.

An diesem Beispiel wird deutlich, daß beim *Adaptive Sampling* einerseits die Zahl der berechneten Strahlen gegenüber dem *Supersampling* mit vergleichbar hoher Auflösung geringer ist, allerdings andererseits die Details C und E weiterhin unberücksichtigt bleiben, sprich es wird wiederum nur eine Verringerung der Aliasing-Effekte erreicht.

2.4 Stochastic Sampling



Die bisher vorgestellten Verfahren haben alle den großen Nachteil Aliasing-Effekte lediglich zu verringern ohne sie vollständig bekämpfen zu können.

Aliasing-Effekte entstehen durch das fassen kontinuierlicher Größen in ein regelmässiges Raster. Benutzt man hingegen ein unregelmässiges Raster, so treten keine Aliasing-Effekte auf.

Man wählt also statt einer festen regelmässigen Matrix von Strahlen beim Supersampling eine stochastische Verteilung innerhalb des Pixels wie in nebenstehender Abbildung.
Aus diesem Grund wird dieses Verfahren als *Stochastic Raytracing* oder auch *Distributed Raytracing* bezeichnet.

Dabei ist darauf zu achten, daß sich die Verteilung der Strahlen benachbarter Pixel unterscheidet, da sonst erneut eine Regelmässigkeit vorliegen und zu Aliasing-Effekten führen würde.

Aliasing-Effekte treten zwar nicht mehr auf, durch die unregelmässige Strahlverteilung tritt nun allerdings Rauschen auf. So bleibt die Problematik, daß z. B. Details, die zwischen Strahlen liegen, verloren gehen, erhalten. Die subjektiv empfundene Bildqualität ist jedoch höher, da Rauschen dem menschlichen Auge deutlich weniger stark auffällt als die regelmässigen Alias-Effekte.

2.5 Statistic Supersampling

Beim *Statistic Supersampling* wird vergleichbar zum *Adaptive Supersampling* Supersampling nur für die Strahlen durchgeführt, für die es nötig erscheint.

Während beim *Adaptive Supersampling* dazu die Strahlen erst berechnet werden, wird beim *Statistic Supersampling* schon vorher mittels statistischer Verfahren die Wahrscheinlichkeit einer Verbesserung der Genauigkeit des Farbwertes bei weiterem Supersampling berechnet. Das eigentliche Supersampling wird dann nur durchgeführt, wenn diese hoch genug ist, man spart also gegenüber dem *Adaptive Supersampling* die Berechnung der Strahlen, bei denen die Differenz der Farbwerte ohnehin gering ist.

3 Beschleunigung

Im Vergleich zu einfacheren Renderingverfahren, die lediglich lokale Beleuchtung verwenden, erzeugt das Raytracing zwar Grafiken mit vergleichsweise hoher Realitätsnähe, allerdings ist der benötigte Rechenaufwand wie in den vorherigen Kapiteln erkennbar wesentlich höher.

In der Praxis ist die Dauer der Berechnung eines Bildes jedoch ein wichtiger Faktor, insbesondere wenn Animationen mit einigen Tausend Einzelbildern erstellt werden. Daher hat man ein starkes Interesse an Verfahren zur Beschleunigung des Raytracings.

3.1 Beschleunigung der Schnittberechnungen

Die Berechnung der Schnittpunkte von Strahlen und Objekten gehört zu den zentralen Elementen eines Raytracers, daher bietet es sich zur Beschleunigung des Raytracings an, diese Schnittberechnungen zu beschleunigen.

Eine offensichtliche Möglichkeit zur Optimierung liegt in der Verwendung effizienter Schnittalgorithmen, hier soll jedoch im folgenden die Benutzung von Bounding-Objekten vorgestellt werden.

3.1.1 Bounding-Objekte

Betrachtet man eine größere Szene, so stellt man fest, daß nur ein geringer Teil der Strahlen ein einzelnes kleines, mitunter aber komplexes Objekt trifft. Dennoch muß für jeden Strahl die relativ aufwendige Schnittberechnung für das komplexe Objekt durchgeführt werden.

Aus diesem Grund benutzt man sogenannte *Bounding-Objekte*, einfache, im Bild nicht sichtbare Objekte mit denen komplexe Objekte umschlossen werden. Schnittests werden zunächst mit dem einfachen Bounding-Objekt durchgeführt, nur wenn ein Strahl dieses schneidet, wird er gegen das umschlossene komplexe Objekt getestet.

Die Verwendung von Bounding-Objekten beschleunigt also negativ[1] ausfallende Schnittests und verlangsamt die positiv[2] ausfallenden. Da jedoch nur ein geringer Teil der Strahlen ein einzelnes Objekt trifft, überwiegt die Beschleunigung.

Bei der Wahl geeigneter Bounding-Objekte gilt es einen guten Kompromiß zwischen möglichst guter Umschließung[3] und möglichst geringer Komplexität des Bounding-

[1]Strahl schneidet das Bounding-Objekt nicht
[2]Strahl schneidet das Bounding-Objekt
[3]Verhältnis der Volumina von Objekt und Bounding-Objekt

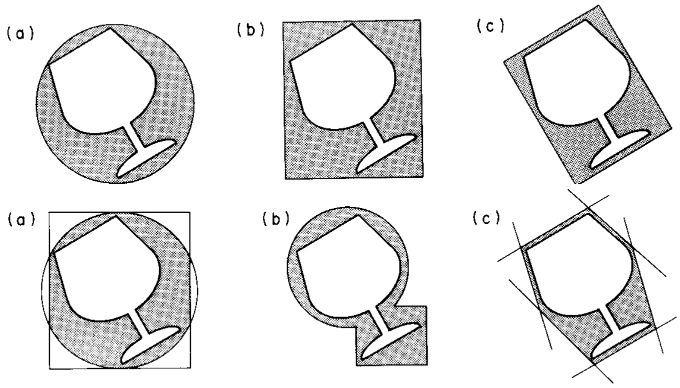


Abbildung 3.1: Verschiedene Bounding-Objekte

Objekts zu finden. Abbildung 3.1 zeigt verschiedene Bounding-Objekte mit von links nach rechts zunehmender Komplexität und besserer Umschließung.

3.2 Reduzierung der Anzahl nötiger Schnittberechnungen

3.2.1 Bounding-Objekt Hierarchien

Bounding-Objekte beschleunigen lediglich einzelne Objekte, nach wie vor wird bei jedem Strahl jedes Objekt der Szene betrachtet. Da also lediglich der Schnittest selbst beschleunigt wird, erreicht man maximal eine Beschleunigung um einen Faktor, also lineare Komplexität bzgl. der Objektzahl.

Da ein Gegenstand einer Szene in aller Regel aus mehreren Primitives besteht[4] und weiterhin oftmals Gruppierungen mehrerer Gegenstände innerhalb einer Szene existieren, bietet sich eine hierarchische Anordnung von Bounding-Objekten an.

Eine solche *Bounding-Objekt Hierarchie* wird als Baum realisiert, bei dem die Wurzel die gesamte Szene, also alle Objekte, umschließt, einzelne Knoten Bounding-Objekte von Objektgruppen darstellen und die Blätter schließlich einzelne Objekte umschließen.

Für jeden Strahl wird nun, beginnend bei der Wurzel, jeweils ein Schnittest mit dem aktuellen Knoten durchgeführt. Ist dieser positiv ausgefallen, so wird rekursiv mit den Nachfolgern fortgefahren, andernfalls die Rekursion für diesen Zweig abgebrochen.

Dadurch verringert sich nun auch die Anzahl der Schnittests, da mit einem einzigen Schnittest ganze Gruppen von Objekten von der weiteren Berechnung ausgeschlossen werden können. Es ergibt sich somit eine logarithmische Komplexität bzgl. der Objektzahl, also eine mit steigenden Objektzahlen deutlich höhere Beschleunigung gegenüber der einfachen Verwendung von Bounding-Objekten.

Die Erstellung der Hierarchie kann auf verschiedene Arten geschehen, z. B. manuell durch den Benutzer oder indirekt durch Übernahme der vom Benutzer festgelegten logischen Objekt-Hierarchie aus dem Szeneneditor. Andere automatische Verfahren beruhen

[4] dies wird insbesondere beim CSG-Modelling deutlich

z. B. auf dem Versickern jedes einzelnen Objektes in der Hierarchie, wobei jeweils der Nachfolger eines Knotens gewählt wird, bei dem das Volumen des zugehörigen Bounding-Objektes am geringsten zunimmt.

3.2.2 Raumunterteilung

Die verschiedenen Techniken der Raumunterteilung, engl. *Spatial Subdivision*, lassen sich zur Beschleunigung beim Raytracing einsetzen, gerade auch in Kombination mit anderen Methoden, wie den Bounding-Objekt Hierarchien. An dieser Stelle soll hierauf jedoch nicht weiter eingegangen werden.

3.2.3 Directional Techniques

Die *Directional Techniques* sind auf der Raumunterteilung basierende Methoden zur Beschleunigung, die die Richtung, in der ein Lichtstrahl verläuft, ausnutzen.

Als Stichworte seien hier nur der *Light buffer*[5], *Ray coherence* und *Ray classification* erwähnt, die Glassner [1] genauer beschreibt.

3.3 Reduzierung der Strahlenzahl

3.3.1 Adaptive Berechnung des Strahlenbaumes

Gewöhnlich wird der Strahlenbaum bis zu einer festen Tiefe berechnet. Bei der adaptiven Berechnung hingegen wird die Berechnung in einem Zweig abgebrochen sobald der Beitrag dieses Zweiges zum Farbwert eine festgelegte Schranke unterschreitet.

3.3.2 Statistic Supersampling

Das *Statistic Supersampling* wurde bereits in Abschnitt 2.5 auf Seite 14 als Verfahren zur Beschleunigung vorgestellt.

3.3.3 Statistische Optimierungen des Anti-Aliasing

Eine vergleichbare Methode ist es, statt der Wahrscheinlichkeit einer Verbesserung bei Berechnung eines bestimmten Strahls, den bereits vorliegenden Farbwert auf statistische Zuverlässigkeit zu untersuchen und das Supersampling bei Erreichen dieser abzubrechen.

3.4 Zusammenfassung von Strahlen

Eine weitere Methode zur Beschleunigung ist es, ähnliche Strahlen zusammenzufassen und diese somit gemeinsam in einem Schritt zu betrachten.

[5] Ein jeder Lichtquelle zugeordneter *Directonal Cube*, der die in der jeweiligen Richtung potentiell angestrahlten Objekte enthält.

3.4.1 Betrachtung von Strahlenbündeln

Dazu betrachtet man statt einzelner unendlich dünner Strahlen z. B. Kegel sowohl gewöhnlicher Art mit runder Grundfläche als auch mit polygonaler, engl. *Beams*, oder Zylinder, engl. *Pencils*.

Durch die Abarbeitung mehrerer Strahlen in einem Schnittberechnungsschnitt erreicht man zwar eine Beschleunigung, allerdings ist die Schnittberechnung selbst deutlich komplizierter, vgl. Abbildung 3.2.

Zudem kommt es zu Einschränkungen bzgl. der verwendbaren Primitive. Zu den Vorteilen zählen hingegen sich ergebendes besseres Anti-Aliasing und optische Effekte, wie z. B. weiche Schatten oder unscharfe Spiegelungen.

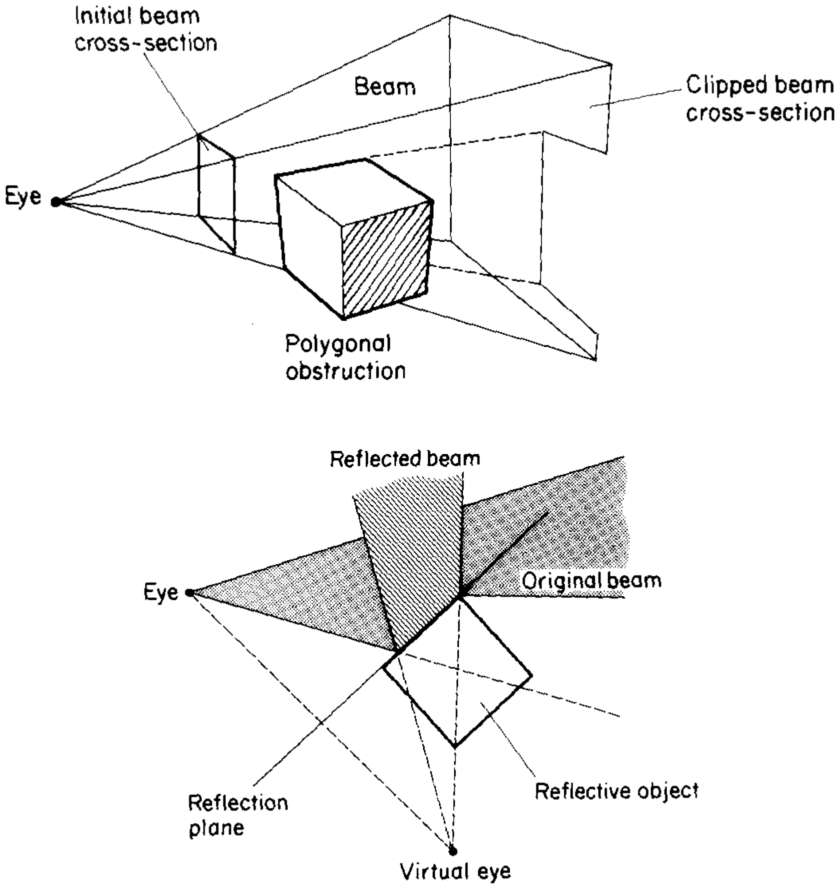


Abbildung 3.2: Schnittberechnung bei Verwendung von *Beams*

Literaturverzeichnis

[1] Andrew S. Glassner: "An introduction to ray tracing", Academic Press, 1989

[2] Gerlach, Grosse: "Physik", B.G. Teubner Stuttgart, 1995

[3] D. Bayer, R. A. Mickelsen, D. Richardson, the POV-Team: "Persistence of Vision Ray-Tracer User's Documentation 3.0", POV-Team (http://www.povray.org/), 1997