|  |  |
| --- | --- |
| Computergrafik | |
| **RayTracing-Algorithmus** | |
| Benotete Abgabe | |
| von | |
| 8978851 | |
| 9482203 | |
| und 9741250 | |
|  | |
| Abgabedatum:10.01.2021 | |
| Dozent | Johannes Riester |
| Bearbeitungszeitraum | 6 Wochen |

Inhaltsverzeichnis

[Abbildungsverzeichnis 3](#_Toc60072597)

[1 Theoretische Prinzipien 4](#_Toc60072598)

[1.1 Grundlagen Raytracing 4](#_Toc60072599)

[1.2 Phong 4](#_Toc60072600)

[1.2.1 Ambient Beleuchtungswert 4](#_Toc60072601)

[1.2.2 4](#_Toc60072602)

[1.2.3 Specular Beleuchtungswert 5](#_Toc60072603)

[1.2.4 Zusammensetzung der Beleuchtungswerte 6](#_Toc60072604)

[2 Algorithmische Umsetzung 7](#_Toc60072605)

[2.1 Kollisionserkennung von Rays mit Objekten 7](#_Toc60072606)

[2.2 Quadratische Lichtquelle 7](#_Toc60072607)

[2.2.1 Berechnung der Schatten 7](#_Toc60072608)

[2.2.2 Berechnung der Beleuchtung 9](#_Toc60072609)

[3 Implementierung 11](#_Toc60072610)

[3.1 Funktionen 11](#_Toc60072611)

[3.2 Klassenmodell 12](#_Toc60072612)

[3.3 Mögliche Optimierungen 13](#_Toc60072613)

[3.3.1 Performance 13](#_Toc60072614)

[3.3.2 Bildgebung 14](#_Toc60072615)

[Quellenverzeichnis 15](#_Toc60072616)

Abbildungsverzeichnis

* [Abbildung 1: Bild mit Umgebungsbeleuchtung 6](#_Toc60146963)
* [Abbildung 2: Bild mit diffuser Beleuchtung 6](#_Toc60146964)
* [Abbildung 3: Bild mit „Specular“-Beleuchtung 7](#_Toc60146965)
* [Abbildung 4: Zusammengesetztes Bild 8](#_Toc60146966)
* [Abbildung 5: Skizze von Rechteck im Raum(auswechseln) 10](#_Toc60146967)
* [Abbildung 6: Bild mit 1 systematischen Shadow Ray 12](#_Toc60146968)
* [Abbildung 7: Bild mit 4 systematischen Shadow Rays 12](#_Toc60146969)
* [Abbildung 8: Bild mit 9 systematischen Shadow Rays 12](#_Toc60146970)
* [Abbildung 9: Bild mit 8 zufälligen Shadow Rays 12](#_Toc60146971)
* [Abbildung 10: Bild mit 9 systematischen und 8 zufälligen Shadow Rays 13](#_Toc60146972)

# Theoretische Prinzipien

## Grundlagen Raytracing

Beim Raytracing handelt es sich um eine Alternative zur Rasterung, der vorherrschenden Methode zum Darstellen von Objekten im Sichtfeld des Betrachters. Dabei wird durch Operationen der Linearen Algebra eine dreidimensionale Szenerie auf einen zweidimensionalen Bildschirm geworfen, wobei die Positionierung des Beobachters im Verhältnis zum Bildschirm und den Objekten in der Szenerie eine Rolle spielt.[[1]](#footnote-1)

Bei dem Raytracing-Verfahren spielt vor allem die Berechnung von Licht, Reflektionen und Schatten eine Rolle. Hier werden vom Beobachter aus Strahlen durch den Bildschirm in die Szenerie geworfen und weiterverfolgt. Abhängig davon, wo diese auftreffen, inwiefern diese reflektieret werden und, wie die Fläche auf die sie auftreffen im Verhältnis zum Licht steht wird dann bestimmt, wie der erste getroffene Punkt beleuchtet wird/welche Farben dieser hat etc. Das Raytracing Verfahren steht dabei konträr zur Realität. Denn tatsächlich treffen Strahlen von Lichtquellen, wie der Sonne oder Lampen zuerst auf Gegenständen auf und fallen dann in unser Auge. Sie gehen also natürlicherweise von der Lichtquelle aus. Beim Raytracing werden allerdings nur die relevanten Strahlen betrachtet, da aus der Perspektive des Betrachters aus nachvollzogen wird, welche Strahlen ins Auge fallen. Das spart Rechenleistung, ist aber offensichtlich auch nicht so realitätsnah, wie die wirkliche Welt.[[2]](#footnote-2)

In der Praxis kann man sich die Realisation von Raytracing grob wie folgt vorstellen:

Aus der Kamera heraus, die eine bestimmte Position im Raum einnimmt, werden Strahlen durch jeden Pixel auf einem Bildschirm geschossen. Dadurch entsteht eine bestimmte Perspektive. Auf der anderen Seite des Bildschirms ist die Szenerie aufgebaut, die man nachvollziehen möchte. Für jeden einzelnen Strahl wird berechnet wo dieser als erstes auftrifft und wie er z.B. potenziell weiterreflektiert wird. Da manche Effekte von Licht aus der Realität viel zu aufwändig zu berechnen sind wird auf Annäherung und Vereinfachungen zurückgegriffen. So z.B. mit dem Phong-Modell, welches in 1.2 näher beschrieben ist. Dort wird z.B. die grundlegende Umgebungsbeleuchtung festgelegt. Aus verschiedenen Zahlenwerten, die für die getroffene Oberfläche und die Lichtwerte gelten und weitere Berechnungen für Reflektion und z.B. das Verhältnis zwischen Lichtquelle und Oberfläche etc. werden einzelne Werte von Farben errechnet, welche auf den anfangs im Bildschirm getroffenen Pixel angewendet werden.

Das Verfahren an sich ist sehr aufwändig, da je nach Pixel sehr viele Strahlen, Kollisionen etc. berechnet werden, vor allem wenn auch noch hochdetaillierte Schatten ins Spiel kommen. Doch mittlerweile ist die Technik der Grafikkarten, mit einer neuen Generation, die auf Raytracing ausgelegt ist, auf der einen und die der Spiele auf der anderen Seite so fortgeschritten dass Raytracing in Echtzeit langsam beginnt auch in die neusten High-End-Grafikspiele mit einzufließen.

## Phong

Beim Phong-Beleuchtungsmodell wird die Beleuchtung eines Punktes im Raum aus den drei Werten, “Ambient”, “Diffuse” und “Specular, zusammengesetzt. Dabei wird nicht das Ziel verfolgt, die Realität möglichst realistisch abzubilden, sondern ein einfaches und schnell zu berechnendes Verfahren zu verwenden.

### Ambient Beleuchtungswert

Der “Umgebungswert” eines Punkts ist nur abhängig von dem Ambient Wert der Oberfläche und dem Ambient Wert der Lichtquelle, bzw. der Lichtquellen. Er stellt somit eine konstante Beleuchtung des gesamten Raums dar, unabhängig von der Position der Lichtquelle oder des Beobachters.

In Abbildung 1 sieht man ein Bild was ausschließlich mit ambienter Beleuchtung erzeugt wurde. Alle Wände sind gleichmäßig beleuchtet.

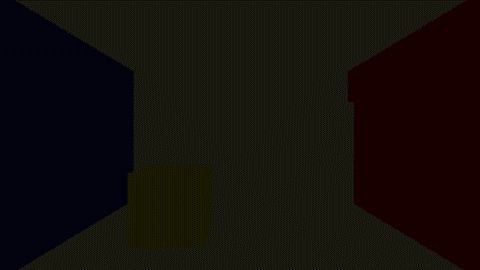


Abbildung : Bild mit Umgebungsbeleuchtung

### Diffuse-Beleuchtungswert

Der diffuse Beleuchtungswert ist abhängig von den „Diffuse“-Werten der Lichtquelle , dem Objekt, sowie dem Winkel der Oberfläche zur Lichtquelle. Steht der Normalenvektor der Oberfläche des Objekts auf dem Vektor zur Lichtquelle ist der diffuse Wert maximal.

In Abbildung 2 ist ein Bild zu sehen, welches nur durch diffuses Licht beleuchtet wurde. Die Flächen sind jetzt unterschiedlich beleuchtet und man sieht auf der dem Licht abgewandten Seite der Quader kein Licht, da das Skalarprodukt aus der Richtung des Lichts und des Normalenvektors der Oberfläche <= 0 ist.

Da der „Ambient“- und der „Diffuse“-Wert nicht von der Position des Betrachters abhängen, kann man diese Werte bei einer Echtzeitanwendung im Voraus berechnen.

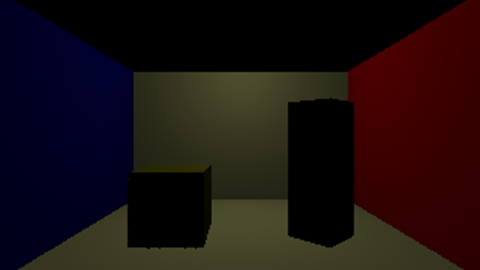


Abbildung : Bild mit diffuser Beleuchtung

### Specular Beleuchtungswert

Der „Specular“-Beleuchtungswert soll spiegelnde, besonders helle Punkte einer Oberfläche betonen, in denen der Winkel zum Licht und zum Betrachter von der Oberflächennormale relativ ähnlich sind. Zur Berechnung wird die optimale Reflektionsachse bestimmt und mit der realen Reflektionsachse (der Oberflächennormale) verglichen.

Je näher sich diese Werte sind (also je größer das Skalarprodukt zwischen den beiden Werten), desto höher ist der Specular-Beleuchtungswert für diesen Punkt. Dieses Skalarprodukt wird dann noch mit einem Wert „shinyness“ potenziert. Durch einen höheren „shinyness“-Wert erhält man eine Fläche die glänzender, bzw. Spiegelnder erscheint, während ein niedriger Wert zu einer eher matten Oberfläche führt.

Abbildung 3 zeigt das Bild nur mit „Specular“-Beleuchtung. Besonders an der Lichtquelle kann man die „harten“ Reflektionen gut sehen.

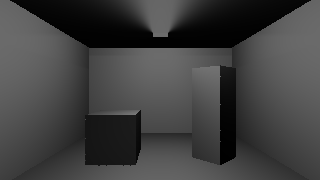


Abbildung : Bild mit „Specular“-Beleuchtung

### Zusammensetzung der Beleuchtungswerte

Für jeden Punkt werden die drei Werte addiert. Dadurch ergibt sich das Bild in Abbildung 4. Durch das Anpassen von den Parametern kann das Bild entsprechend verändert werden.[[3]](#footnote-3)

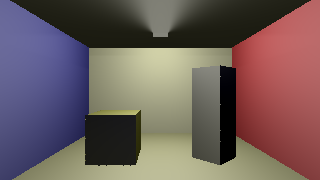


Abbildung : Zusammengesetztes Bild

# Algorithmische Umsetzung

## Kollisionserkennung von Rays mit Objekten

Bei einem der zentralsten Funktionen in einem Raytracing-Algorithmus handelt es sich um die Kollisionserkennung. Diese ist notwendig um bei den verschiedenen Strahlen deren Weg nachvollzogen werden soll, nachzuvollziehen, wo diese als erstes auftreffen. Eine Grundlage für die Anwendung einer solchen Funktion ist das Verständnis von linearer Algebra und das Verständnis für den dreidimensionalen Raum und vor allem die Implementierung der einzelnen Konzepte, wie die eines Strahles oder die Darstellung von Objekten im Raum. Näher wird hierauf in Abschnitt 3.1 eingegangen. An dieser Stelle wird auf die Implementierung von Strahlen und die von Objekten bzw. Flächen eingegangen.

Ein Strahl besteht immer aus einem Ursprung und einem Richtungsvektor. Wandert man vom Ursprung aus entlang des Richtungsvektors erhält man alle Punkt, die vom Strahl erreicht werden können. Da der Richtungsvektor eben nur die Richtung anzeigt sollte er normalisiert werden. Das bedeutet, dass die Länge des Vektors eins ergibt, er aber trotzdem weiterhin in die gleiche Richtung zeigt. So lässt sich auch später die Distanz, die der Strahl zurücklegen muss, ganz einfach und ohne weitere Berechnung bestimmen.

Je nach Verwendungen von dazustellenden Objekten müssen auch Möglichkeiten implementiert werden, Schnittpunkte mit diesen auszurechnen. In unserem Beispiel lassen sich alle Objekte durch Rechtecke darstellen. Egal ob Wände des Raums oder die einzelnen Flächen eines Würfels. Alle Objekte sind eigentlich durch zweidimensionale Flächen im Raum implementiert. Bei einem Würfel liegen dann sechs einzeln definierte Flächen im Raum, die aneinander anschließen. Vereint sind sie durch einzelne Eigenschaften, die für das gesamte Objekt gelten, vgl. Phong Modell aus Abschnitt 1.2. Die einzelnen Flächen wiederrum sind definiert durch einen Stützvektor, also einen Punkt im Raum und zwei direktionale Vektoren, die die Fläche aufspannen. Dabei haben in diesem Fall die direktionalen Vektoren eine bestimmte Länge. Diese entspricht der Seitenlänge des Rechtecks. Verbindet man jetzt die zwei direktionalen Vektoren miteinander, die vom Stützvektor ausgehen, indem man für die jeweiligen Vektoren auf der Spitze des anderen direktionalen Vektors parallel den gleichen Vektor zeichnet erhält man eine rechteckige Fläche, wie zu sehen.

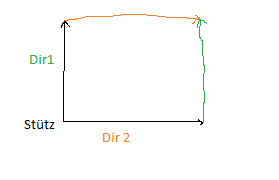


Abbildung : Skizze von Rechteck im Raum(auswechseln)

Nun soll anhand dieser Gegebenheiten im Raum errechnet werden, wo mögliche Schnittpunkte zwischen Strahlen und den Flächen liegen und welcher der möglichen Schnittpunkte der erste ist, also welche Fläche als erstes geschnitten wird.

Hierzu ist ein lineares Gleichungssystem notwendig.

Mit dem obigen Gleichungssystem lassen sich die einzelnen Koordinatenteile, also x/y/z senkrecht und jeweils für die Fläche links und den Strahl rechts berechnen. So erhält man ein lineares Gleichungssystem dritter Ordnung. X steht hierbei für die Ausdehnung des ersten direktionalen Vektors und Y für die des zweiten. Mit Y lässt sich die Ausdehnung des normalisierten Vektors des Strahls berechnen. Ist das Gleichungssystem lösbar so gibt es einen Schnittpunkt. Mit X und Y und dem Rest der Flächengleichung könnte man den Schnittpunkt errechnen. Aber auch mit Z und dem Rest der Gleichung für einen Strahl lässt sich dieser Punkt errechnen. Eine Besonderheit hier: Aufgrund der Normalisierung des Richtungsvektor des Strahls gibt Z zudem die Distanz zwischen Ursprung des Strahls und der Fläche an.

Wie genau diese Formel implementiert wurde und zu welchen Problemen die Kollisionserkennung führt, ist näher in Abschnitt 2.1 beschrieben.

## Quadratische Lichtquelle

Da eine quadratische Lichtquelle gefordert ist, aber die in Kapitel 1 beschriebene Phong-Methode nur mit eindimensionalen Lichtquellen arbeitet, stellt das Berechnen der Beleuchtung eines Punktes ein besonderes Problem dar. Zusätzlich ist der einfache Ansatz zur Schattenbestimmung eines Punktes durch Werfen eines Strahls zur (punktförmigen) Lichtquelle, bedingt durch die zweidimensionale Ausdehnung der Lichtquelle, nicht möglich. Um weiche Schatten zu erhalten muss also ein anderer Ansatz gewählt werden.

### Berechnung der Schatten

Zur Berechnung der Schatten wird ein primitiver Algorithmus verwendet, der die Schattenbestimmung von einer eindimensionalen Lichtquelle auf eine mehrdimensionale Lichtquelle portiert. Statt zu überprüfen, ob zwischen zwei Punkten ein Objekt liegt, wird überprüft wie viele Strahlen ungestört von einem Punkt auf dem Objekt zu mehreren anderen Punkten auf dem Licht reichen. Es werden also von dem zu prüfenden Punkt mehrere „Schattenfühler“ (Shadow Rays) zur Lichtquelle geschickt und aus dem Anteil der angekommen Fühler ein Schattenwert für diesen Punkt berechnet.

TODO: Skizze mit Zeichentablet einfügen

Diese Methode approximiert die Lichtfläche nur durch mehrere Punkte und stellt daher kein perfektes Ergebnis dar und nur stufenförmige Schatten (siehe Abbildung 6). Die Punkte, durch die die Lichtquelle repräsentiert wird, können dabei systematisch(siehe Abbildung 5 bis Abbildung 7), zufällig(siehe Abbildung 8) oder gemischt(siehe Abbildung 9) gewählt werden. Für den finalen Render wurden jedoch nur systematische Shadow Rays benutzt, da diese für ein schöneres Bild sorgen und die ungenutzten zufälligen Shadow Rays später zu Fehlern im Bild geführt haben (Siehe Abbildung 8 und Abbildung 9).

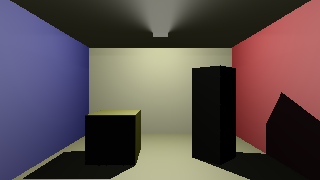


Abbildung : Bild mit 1 systematischen Shadow Ray

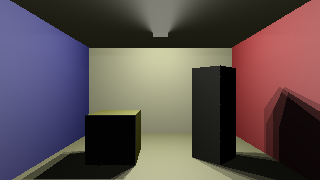


Abbildung : Bild mit 4 systematischen Shadow Rays

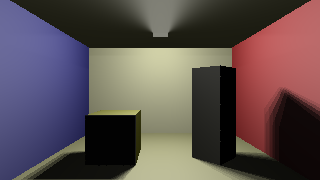


Abbildung : Bild mit 9 systematischen Shadow Rays

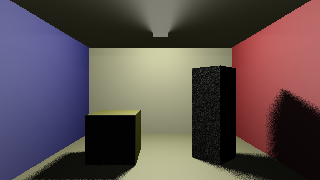


Abbildung : Bild mit 8 zufälligen Shadow Rays

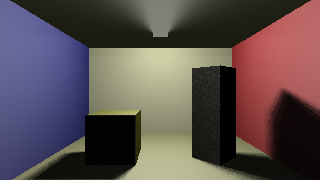


Abbildung : Bild mit 9 systematischen und 8 zufälligen Shadow Rays

### Berechnung der Beleuchtung

Bei der Berechnung der Beleuchtung für eine quadratische Lichtquelle nach dem Phong-Modell ist nur der „Diffuse“- und „Specular“-Wert wichtig, da nur diese von der Position des Lichts abhängen. Für die Berechnung dieser Werte wird daher der für den Punkt auf dem Objekt der optimale Punkt auf der Lichtfläche zur Berechnung verwendet.

Bei der Berechnung des Specular-Wertes wird anhand der realen Reflektionsachse (dem Normalenvektor der Oberfläche) ein reflektierter Ray von der Kamera aus gebildet. Wie weit man diesem Ray folgen muss um auf der Lichtebene zu landen wird mit folgender Formel berechnet:

Wenn der Ray parallel zur der Lichtebene ist () wird zu der Richtung des Rays ein kleiner Offset hinzugefügt, um den Schnittpunkt nahe Unendlich zu legen. Der herausgefundene Schnittpunkt wird anschließend durch eine Basisänderung so vom dreidimensionalen in den zweidimensionalen Raum projiziert, so dass ein Wert im Bereich einem Punkt auf der Lichtebene innerhalb des Lichts entspricht (Für die Transformation wird die Pseudoinverse einer Matrix, die aus den aufspannenden Vektoren des Lichts besteht, benutzt). Ist n positiv wird der zu dem Schnittpunkt nächste Wert im Bereich rücktransformiert, wenn n negativ ist, wird der am weitesten entfernteste Punkt rücktransformiert. Dieser rücktransformierte, nun wieder im dreidimensionalen Raum liegende, Punkt wird dann für die Berechnung der optimalen Lichtachse verwendet.

Die Berechnung für die optimale Position des Lichts für den diffusen Wert verläuft ähnlich, nur wird hier als initialer Ray nicht der reflektierte Vektor von der Kamera verwendet, sondern der Normalenvektor der Oberfläche.

//TODO: ganz viele skizzen einfügen

# Implementierung

## Implementierung des 3-dimensionalen Raums

## Funktionen

An dieser Stelle sollen interessante Funktionen der einzelnen Klassen beleuchtet werden. Ein Überblick über alle Klassen und deren Funktionen wird dabei in Abschnitt 3.3 gegeben.

### Funktion zur Schnittpunktberechnung

Wie schon in Abschnitt 2.1 beschrieben ist für die Implantierung des Raytracers eine Kollisionserkennung notwendig. Im Programm existiert eine Klasse „surfaces“ diese wird genutzt um alle Oberflächen, die im Modell existieren zusammenzufassen unter einem Objekt. Auf diese Oberflächen kann dann eine Klassenfunktion der Klasse „surfaces“ mit dem Namen „getCollisionObject“ angewandt werden. Die Funktion erwartet einen „ray“, also einen Strahl. Dieser verfügt auch in der implementierten Klasse nur über die Eigenschaften „origin“, also Ursprung und die Eigenschaft „normDirection“, also den normalisierten Richtungsvektor des Strahls.

Die Funktion aus Abschnitt 2.1 lässt sich umstellen, sodass eine Seite abhängig ist von den Koeffizienten X/Y/Z und eine Seite einen Konstanten Wert hat.

So lassen sich im Programm die folgenden Matrizen bilden:

Die Werte werden aus dem „ray“-Objekt und einer einzelnen Fläche, die Teile einer Liste, eines „surfaces“-Objekts ist, herausgelesen. Angenommen man überprüft nur eine Fläche auf einen Schnittpunkt so wird versucht das Gleichungssystem zu lösen. Dazu wird die linalg.solve-Funktion genutzt von NumPy genutzt. Diese empfängt als Parameter die 3x3 Koeffizientenmatrix und die 3x1 Ergebnismatrix. Bei nicht Lösbarkeit wird eine Exception geworfen, die abgefangen werden kann. Bei Lösbarkeit wird als Rückgabewert die Variablenmatrix mit X/Y/Z zurückgegeben.

Gibt es eine Lösung so wird überprüft ob X und Y größer als 1 oder kleiner als 0 sind. Da es sich nur dabei um tatsächliche Punkte auf der von uns gewollten Fläche handelt. Auch wird überprüft ob Z positiv ist, sodass der Strahl nicht in die falsche Richtung geworfen wird.

Da wir nun aber mehrere Flächen haben und es möglich ist, dass ein Strahl Schnittpunkte mit mehreren dieser Flächen hat, muss auch noch überprüft werden, welche Fläche als erstes getroffen wird. Dazu wird die erste Fläche die einen Schnittpunkt hat, zusammen mit der Variablenmatrix zwischengespeichert und später der Z-Wert der Variablenmatrix mit dem Z-Wert des neuen Schnittpunktes mit einer anderen Fläche verglichen. Ist das neue Z kleiner so wird die neue Fläche zusammen mit der neuen Variablenmatrix zwischengespeichert. Die Fläche, die am Ende übrig bleibt, wird zusammen mit der Strecke, die der Strahl zurücklegt, also Z, zurückgegeben.

### Probleme mit Schnittpunktberechnung

## Klassenmodell

## Mögliche Optimierungen

Neben der leichten Erweiterbarkeit des Programms durch seinen objektorientierten Aufbau, kann man das Programm auch noch abwandeln, um eine schnellere Berechnung zu erreichen. Zusätzlich kann man das bilderzeugende Verfahren anpassen, um realistischere oder interessantere Bilder zu bekommen.

### Performance

Das implementierte Programm ist denkbar ineffizient. So braucht die Berechnung eines Bildes in Full-HD (1080\*1920px) mit 64 Shadow Rays und einer maximalen Raytracing-Tiefe von 4 (also maximal 256 Shadow Rays pro Pixel) etwa 15 Stunden auf einer modernen CPU. Dies liegt daran, dass Python als interpretierte Sprache langsamer ist als eine Sprache die als Maschinencode ausgeführt wird, wie etwa C.

Trotzdem kann man durch konsequentes Nutzen von in C implementieren Modulen von Python, wie etwa Numpy, und das Verwenden von in C umgesetzten builtins von Python, wie der map() Befehl. Durch die Ersetzung von einigen Schleifen durch map() konnte die Laufzeit um fast ein Zehntel reduziert werden.

Neben der effizienteren Ausnutzung der vorhandenen Ressourcen kann man die vorhandenen Ressourcen erhöhen, indem man mehr Prozessorkerne zur Berechnung einsetzt. Durch den Einsatz des Python Moduls „multiprocessing“ (Standard Python Modul) konnte die Berechnungszeit auf einem System mit vier Kernen ca. auf ein Drittel reduziert werden.

|  |  |
| --- | --- |
| Berechnungsverfahren | Zeit |
| Single Process | ~15:00 |
| Multi Process | ~6:20 |

Tabelle : Vergleich der Rechenzeiten eines 480\*853px Bildes mit 16 Shadow Rays

*Da multiprocessing kein erlaubtes Modul ist, gibt es im Github einen extra Branch der dieses Modul verwendet.*

### Bildgebung

Quellenverzeichnis

(kein Datum). Von https://medium.com/swlh/ray-tracing-from-scratch-in-python-41670e6a96f9 abgerufen

Phong, B. T. (1975). *Illumination for Computer Generated Pictures.*

Wegscheider, A. (2009/2010). *Facharbeiten: Technische-Universität-München-Website.* Abgerufen am 28. 12 2020 von Technische-Universität-München-Website: https://www.edu.tum.de/fileadmin/tuedz01/www/Sch%C3%BClerkonferenz/Facharbeiten\_2010/andreas\_wegscheider\_2010.pdf

1. (Wegscheider, 2009/2010, S. 3) [↑](#footnote-ref-1)
2. (Wegscheider, 2009/2010, S. 3-5) [↑](#footnote-ref-2)
3. (Phong, 1975); (Aflak, 2020) [↑](#footnote-ref-3)