



**School of
Engineering**

InIT Institut für angewandte
Informationstechnologie

Bachelorarbeit in Informatik

Level of Detail Generierungs-System für 3D-Webapplikationen (max. 4 Zeilen)

Autoren

Marc Berli
Simon Stucki

Hauptbetreuung

Dr. Gerrit Burkert

Nebenbetreuung

Titel Vorname Name

Industriepartner

Externe Betreuung

Datum

17.04.2021

Bitte füllen Sie das Titelblatt aus und berücksichtigen Sie Folgendes:

- ➔ Bitte auf keinen Fall Schriftart und Schriftgrösse ändern. Text soll lediglich überschrieben werden!
- ➔ Bitte pro Tabellenzeile max. 4 Textzeilen!
- Vorlage: Haben Sie die richtige Vorlage gewählt? → Logo Institut/Zentrum
- Titel: Fügen Sie Ihren Studiengang direkt nach dem Wort „Bachelorarbeit“ ein (max. 2 Zeilen).
- Titel der Arbeit: Überschreiben Sie den Lauftext mit dem Titel Ihrer Arbeit (max. 4 Zeilen).
- Autoren: Tragen Sie Ihre Vor- und Nachnamen ein (bitte alphabetisch nach Name).
- Betreuer: Tragen Sie Ihren Betreuer / Ihre Betreuer ein (bitte alphabetisch nach Name).
- Nebenbetreuung: Falls Sie keine Nebenbetreuung haben → bitte ganze Tabellenzeile löschen.
- Industriepartner: Falls Sie keinen Industriepartner haben → bitte ganze Tabellenzeile löschen.
- Externe Betreuung: Falls Sie keine ext. Betreuung haben → bitte ganze Tabellenzeile löschen.
- Datum: Bitte aktuelles Datum eintragen.
- Schluss: Am Schluss löschen Sie bitte den ganzen Beschrieb (grau) und speichern das Dokument als pdf. ab.

Erklärung betreffend das selbständige Verfassen einer Bachelorarbeit an der School of Engineering

Mit der Abgabe dieser Bachelorarbeit versichert der/die Studierende, dass er/sie die Arbeit selbständig und ohne fremde Hilfe verfasst hat. (Bei Gruppenarbeiten gelten die Leistungen der übrigen Gruppenmitglieder nicht als fremde Hilfe.)

Der/die unterzeichnende Studierende erklärt, dass alle zitierten Quellen (auch Internetseiten) im Text oder Anhang korrekt nachgewiesen sind, d.h. dass die Bachelorarbeit keine Plagiate enthält, also keine Teile, die teilweise oder vollständig aus einem fremden Text oder einer fremden Arbeit unter Vorgabe der eigenen Urheberschaft bzw. ohne Quellenangabe übernommen worden sind.

Bei Verfehlungen aller Art treten die Paragraphen 39 und 40 (Unredlichkeit und Verfahren bei Unredlichkeit) der ZHAW Prüfungsordnung sowie die Bestimmungen der Disziplarmassnahmen der Hochschulordnung in Kraft.

Ort, Datum:

.....

Unterschriften:

.....

.....

.....

Das Original dieses Formulars ist bei der ZHAW-Version aller abgegebenen Bachelorarbeiten zu Beginn der Dokumentation nach dem Titelblatt mit Original-Unterschriften und -Datum (keine Kopie) einzufügen.

Zusammenfassung

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Vorwort

Die Autoren der Arbeit sind seit mehreren Jahren in der Webentwicklung tätig. Der rapide Fortschritt in der Webentwicklung reicht insbesondere vom Engagement der vielen Freiwilligen her. Insbesondere für 3D Visualisierungen sind in den vergangenen Jahren viele neue Möglichkeiten entstanden. Aufgrund des Alters sind die Tools noch nicht auf dem Level der anderen Plattformen. Diese Arbeit stellt insofern einen Versuch dar, die Community zu unterstützen und die Möglichkeiten im Web zu erweitern.

Inhaltsverzeichnis

1. Einleitung	6
1.1. Ausgangslage	6
1.2. Zielsetzung	6
2. Grundlagen	7
2.1. Kontext	7
2.2. 3D-Rendering im Web	7
2.3. 3D-Modelle	8
2.4. Transformation von Modellen	11
2.5. Graphics Pipeline	14
2.6. Performanzoptimierung	15
2.7. Einführung Level of Detail	16
2.7.1. Ansätze für LOD Artefakte	17
2.7.2. Vergleich Algorithmen	18
3. Vorgehen	19
3.1. Nutzen LOD	19
3.1.1. Definition Benchmark	19
3.2. Vergleich LOD Systeme	21
3.3. Surface Simplification Algorithmus	22
3.3.1. Fehler Metrik	22
3.4. Pipeline Integration	22
3.5. Automatische Generierung von Detail Levels	23
3.6. Levelwahl während Laufzeit	23
4. Resultate	24
4.1. Benchmark	24
4.2. LOD Pipeline	24
5. Diskussion und Ausblick	25
6. Verzeichnisse	26
A. Anhang	31
A.1. Aufgabenstellung	32
A.2. Anhang 2	33

1. Einleitung

1.1. Ausgangslage

Dank der Rechenleistung auf modernen Geräten ist es möglich anspruchsvolle 3D Visualisierungen auf diversen Geräten anzuwenden. Da diese Applikationen sehr rechenintensiv sind und gerade mobile Geräte in ihrer Rechenleistung beschränkt sind, ist Performance Optimierung im 3D-Rendering unabdinglich. Insbesondere die Komplexität der Modelle hat einen signifikanten Einfluss auf die Leistung. Eine Möglichkeit zur Optimierung ist das Anzeigen von vereinfachten Modellen ab bestimmten Distanzen zum Betrachter. So kann z.B. ein Modell in grosser Distanz vereinfacht dargestellt werden, solange bei genauer Betrachtung mehr Details sichtbar werden. In diversen Rendering Engines ¹ gibt es deshalb Möglichkeiten für das Verwenden von sogenannten Level Of Details (LOD) Artefakten. Für Game Engines ² wie Unreal oder Unity gibt es bewährte Möglichkeiten, um den Einsatz von LOD Artefakten zu vereinfachen. Zurzeit gibt es im Web Bereich keine weit verbreitete Möglichkeit für das Generieren von LOD Artefakten.

1.2. Zielsetzung

Ziel der Arbeit ist es, ein Tool zu entwickeln, das den Umgang mit LOD Artefakten im Web vereinfacht. Hierfür muss vorab der Beleg erbracht werden, dass das Laufzeitverhalten der Applikation mit LOD Artefakten verbessert werden kann. Des Weiteren muss ein Algorithmus entwickelt werden, welcher es erlaubt Modelle zu vereinfachen ohne die grobe geometrische Form zu verlieren. Im Anschluss muss das Tool zur Verfügung gestellt werden, sodass es in der Praxis eingesetzt werden kann. Diese Zielsetzung definiert somit die verschiedenen Schritte der Arbeit.

¹Teilprogramm, das zuständig für die Darstellung von Grafiken ist

²Framework, das für den Spielverlauf und dessen Darstellung verantwortlich ist

2. Grundlagen

2.1. Kontext

In der Arbeit wird der Fokus auf die Web-Plattform gelegt. Die meisten Grundlagen sind jedoch auch unabhängig davon zutreffend. Für die Vereinfachung wird im Folgenden der Begriff "Das Web" für die Plattform von verschiedenen Web Technologien genutzt. Dies beinhaltet insbesondere vom W3C veröffentlichte Spezifikationen.

2.2. 3D-Rendering im Web

3D-Visualisierungen werden in vielen Branchen verwendet. Virtual Reality, CAD Anwendungen oder Computerspiele sind bekannte Anwendungsgebiete, welche 3D Visualisierungen einsetzen. Dank leistungstärkeren Geräten sind in vielen weiteren Bereichen realere Visualisierungen möglich.

Anwendungen können auf spezifische Hardware wie zum Beispiel die Spielkonsolen PlayStation oder Xbox ausgerichtet sein. Dies hat den Nachteil, dass die Verteilung der Software aufwändiger ist, da spezifische Hardware notwendig ist. Für gewisse Anwendungsgebiete ist dies kein Problem.

Für hardwareunabhängige Anwendungen eignet sich das Web hervorragend. Viele Benutzer haben Zugang zu einem Desktop, Tablet oder Mobiltelefon. Somit ermöglicht das Webanwendungen mit weniger Aufwand einem grossen Zielpublikum zugänglich zu machen. Seit einigen Jahren ist es auch möglich, 3D Visualisierungen im Web zu realisieren. Als Basis dafür dient meist das von der Khronos Group entwickelte WebGL, das von allen modernen Browsern unterstützt wird. WebGL ist eine low-level JavaScript API für 3D Visualisierungen. [1] Alternativ zu WebGL wird zurzeit ein weiterer Standard entwickelt: WebGPU. Dieser ist zur Zeit des Schreibens noch in Entwicklung und wird deshalb nicht weiter berücksichtigt, auch wenn ein grosses Potenzial vorhanden ist. [2] Die Unabhängigkeit der Hardware bedeutet jedoch auch, dass Optimierung der Performance in Webanwendungen unabdinglich ist, um allen Benutzern ein optimales Erlebnis zu ermöglichen. Im Vergleich zu fixen Hardware Anwendungen ist es realistisch, dass eine Webanwendung sowohl auf einem leistungsfähigen Desktop Computer als auch auf einem günstigen Mobilgerät verwendet wird.

Zudem ist WebGL eine junge Technologie und wurde erst 2011 veröffentlicht – verglichen mit dem initialen Release Date von OpenGL¹ welches im Jahre 1992 publiziert wurde. [1, 3] Nicht nur das Alter, sondern auch die Natur der Web Plattform haben dazu beigetragen, dass WebGL ein langsames Wachstum verspürt hat. Um einen Webstandard wie

¹Spezifikation einer Programmierschnittstelle zur Entwicklung von 2D und 3D-Grafikanwendungen

WebGL einsetzen zu können müssen alle grossen Browser die Spezifikation implementieren. Ansonsten kann der grosse Vorteil des Webs – einfache Verteilung an alle Benutzer – nicht in vollem Umfang genutzt werden. So hat zum Beispiel Internet Explorer 10 keinen Support und es wurde erstmals Ende 2013 möglich im Internet Explorer 11 3D Anwendungen für ein breites Publikum zu entwickeln.

2.3. 3D-Modelle

Ein Modell stellt ein Objekt aus der realen Welt vereinfacht dar. 3D-Modelle können vereinfacht als Gruppe von Punkten definiert werden. Um im dreidimensionalen Raum Objekte visualisieren zu können sind mindestens drei Punkte notwendig. Punkte von 3D-Modellen werden im folgenden als Vertex (Eckpunkte) bezeichnet und können als Vektor definiert werden. So können wir einen Vertex am Ursprung eines Koordinatensystems definieren als:

$$V = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Eine Sammlung von drei *Vertices* bildet ein *Triangle* (Dreieck). Hiermit kann ein *Triangle* generell wie folgt definiert werden:

$$T = \begin{bmatrix} V \\ V \\ V \end{bmatrix}$$

Um komplexere Formen wie Vierecke (*Quads*) zu bilden, werden jeweils mehrere *Triangle* kombiniert. Abbildung 2.1 zeigt die Vereinfachung eines *Quads* in zwei *Triangle*.



(a) Polygon

(b) Polygon als Sammlung von Triangles

Abbildung 2.1.: Vereinfachung von Polygone als Sammlung von *Triangles*

Für eine Sammlung von Punkten wird generell der Term Polygon verwendet. Ein Modell besteht aus einer beliebigen Anzahl Polygone.

Die Verbindung zwischen zwei Vertices ist eine sogenannte Edge (Kante). Verbindet man die Punkte eines Polygons und füllt die Fläche, ergibt sich schlussendlich ein Face (Fläche).

Weitere Attribute

Neben den geometrischen Attributen verfügt ein Modell ebenfalls über visuelle Attribute. So wird für jeden Vertex ein sogenannter *Normal* definiert. Ein *Normal* ist ein Vektor der im einfachsten Fall senkrecht zu den zwei an diesem Vertex verbundenen Edges verläuft. In der Geometrie ist dies auch bekannt als Normale. *Normals* werden häufig für das Berechnen von Reflexionen verwendet. *Normals* werden auch für gewisse Performanz Optimierungen eingesetzt, dazu mehr in Absatz 2.6.

Um die Oberfläche von Modellen zu definieren wird häufig ein sogenanntes Texture Mapping² durchgeführt. Hierfür sind zusätzliche Informationen für ein Modell notwendig. In der Praxis finden viele weitere Methoden Anwendung, auf diese wird jedoch hier nicht weiter eingegangen.

Formate

Um ein 3D Modell in einer Anwendung zu verwenden, muss ein entsprechendes Format verwendet werden. Hierfür steht eine Vielzahl von Optionen zur Auswahl. Aufgrund der Menge wird hier jedoch nur oberflächlich auf die bekanntesten Formate eingegangen.

OBJ Wavefront OBJ ist ein offenes Dateiformat das von Wavefront Technologies 1989 entwickelt wurde. Das Format ist jedoch speichertechnisch ineffizient und verfügt zudem nur über ein limitiertes Feature Set. Des Weiteren gibt es keine zentrale Instanz, welche eine Spezifikation liefert und Informationen sind deshalb schwerer zu finden. [4]

FBX FBX ist ein proprietäres Format, welches von Autodesk verwaltet wird. Das Verwenden von FBX Daten ist jedoch offiziell nur mit einer C++ FBX SDK möglich, welche für das Web nicht geeignet ist. Aufgrund der proprietären Natur gibt es keine Bestrebungen das Format offener zu gestalten.

glTF Seit 2015 gibt es ein offenes, modernes sowie auf optimale Speichernutzung fokussiertes Format: glTF. Dieses Format wird von der Khronos Gruppe entwickelt. [5] In dieser Arbeit wird es als Austauschformat verwendet, deshalb wird im folgenden genauer auf das Format eingegangen.

Die meisten 3D-Grafikprogramme wie Blender (.blend) verwenden ihre eigenen Dateiformate. Diese Dateiformate sind jedoch für den Einsatz in einer Applikation ungeeignet da sie insbesondere nicht für die Laufzeitperformanz optimiert sind. Deswegen brauchte

²Verfahren, mittels 2D-Bildern 3D-Objekte zu gestalten

es für jedes Eingangsformat für jedes Ausgangsformat einen *converter*, wie in Abbildung 2.2 ersichtlich.

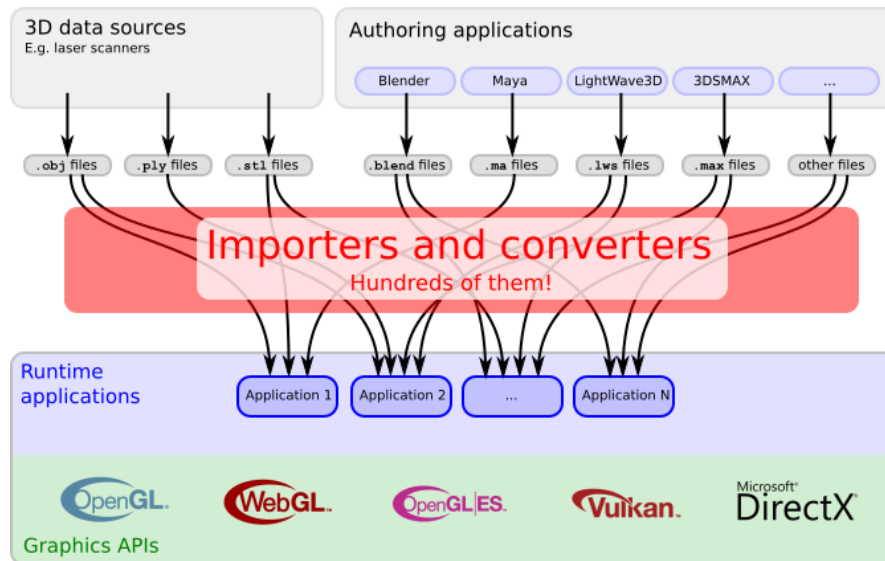


Abbildung 2.2.: Convertierungspipeline vor glTF [5]

Durch die immer breiter werdende Nachfrage nach 3D-Applikationen wurde ein Format benötigt, dass zum einen Applikationsunabhängig verwendet und zum anderen performant im Web eingesetzt werden kann. [5] Durch diese Abstraktion kann eine Vielzahl *converter* vermieden werden und nur wenige, sehr spezifische Dateiformate benötigen noch einen einzigen *converter*. Diese neue Pipeline ist in Abbildung 2.3 dargestellt. [5]

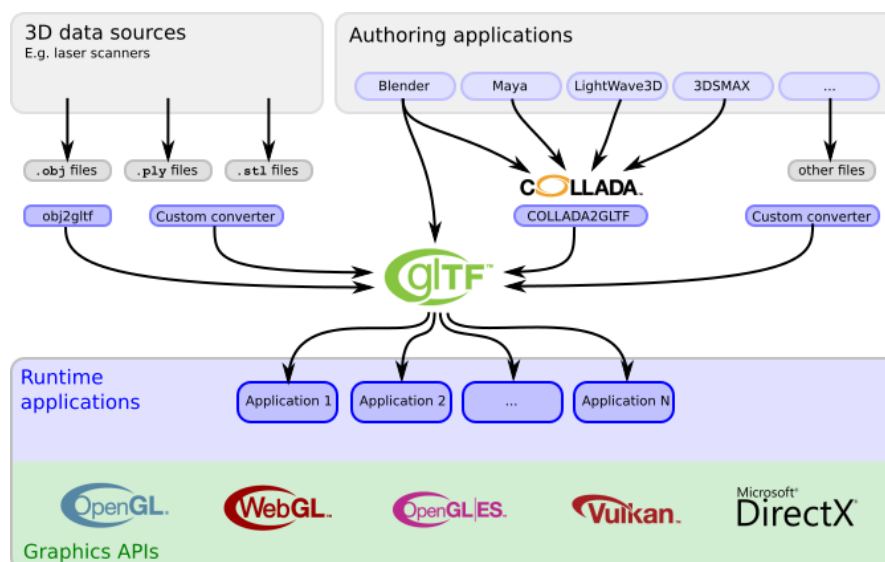


Abbildung 2.3.: Convertierungspipeline mit glTF [5]

Die Basisstruktur von glTF basiert auf JSON. Darin ist die komplette Szenerie des Modells beschrieben und in Abbildung 2.4 aufgezeigt.[5]



Abbildung 2.4.: glTF Datenstruktur [5]

Die wichtigsten Elemente dieser Struktur kurz erläutert:

- **Scene:** Einstiegspunkt der Szenerie und verweist auf alle Top-Level *nodes*.
- **Node:** Kann eine Transformation (Rotation, Translation oder Skalierung) beinhalten oder eine Kamera, *Skin*, Animation, ein *Mesh* oder *Childnodes* referenzieren.
- **Mesh:** Beschreibt ein geometrisches Objekt und verweist auf *accessor*, welcher die effektiven geometrischen Daten beinhaltet und auf *material*, das beschreibt, wie das Objekt beim Rendern aussehen soll.
- **Accessor:** Verweist auf die Binärdaten (*BufferView*), welche die effektiven Eigenschaften für *meshes*, *skins* und *animations* kompakt beinhaltet.
- **BufferView:** Definiert eine Ansicht (Länge, Ort, Typ) auf einen *Buffer*.
- **Buffer:** Verweist auf einen Block von Binärdaten, welchen die effektiven Daten des 3D-Modells platzsparend beinhaltet.
- Weitere Elemente, welche im Rahmen dieser Arbeit nicht relevant sind und nicht im Detail betrachtet werden, sind: *camera*, *skin*, *material*, *animation*, *texture*.

2.4. Transformation von Modellen

Um ein Modell vereinfachen zu können, muss es verändert werden. Diese Veränderungen können in Basis Operationen vereinfacht erläutert werden.

Im Folgenden werden Transformationen mithilfe einer 2D Visualisierung erläutert. Das Modell kann jedoch ebenfalls im dreidimensionalen Raum sein – die Funktionsweise bleibt identisch.

Für die folgenden Beispiele wird jeweils das Modell aus Abbildung 2.5 verwendet.

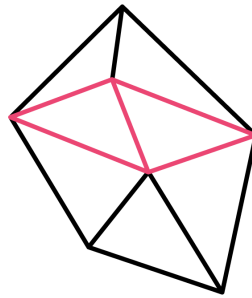
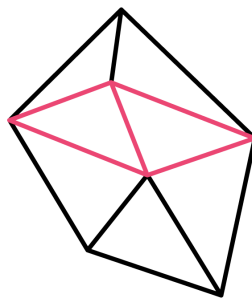


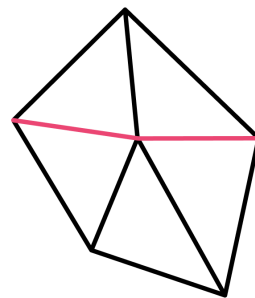
Abbildung 2.5.: Modell Basis

Edge Collapse

Hierbei werden zwei nebeneinanderliegende *Vertices* kombiniert. Durch diese Operation kann ein *Vertex* entfernt werden. In Abbildung 2.6 kann die Anzahl *Triangles* um zwei verringert werden. Beim *Edge Collapse* wird ein neuer *Vertex* definiert und zwei bestehende entfernt. Die Anzahl entfernter *Triangles* ist abhängig von der Situation. Unter Umständen kann so auch nur ein einzelner *Triangle* entfernt werden. Die Umkehrfunktion nennt man Vertex Split.



(a) Modell Ausgangslage

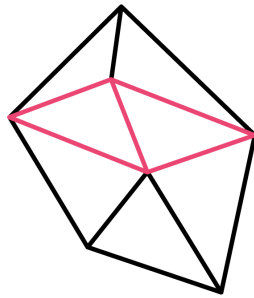


(b) Edge Collapse

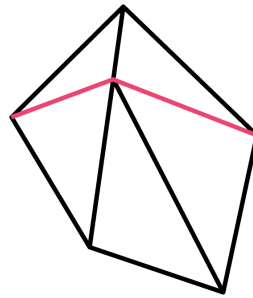
Abbildung 2.6.: Transformation mittels Edge Collapse

Halfedge Collapse

Hierbei wird ein Vertex direkt entfernt und alle Edges auf einen danebenliegenden Vertex zusammengelegt. Bei dieser Operation muss kein neuer *Vertex* definiert werden sondern ein bereits bestehender *Vertex* kann wiederverwendet werden. Wie in Abbildung 2.7 ersichtlich können auch in diesem Fall zwei *Triangles* entfernt werden.



(a) Modell Ausgangslage

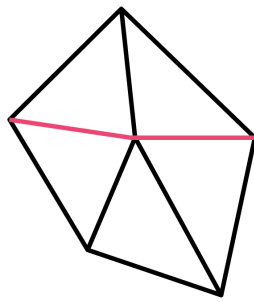


(b) Halfedge Collapse

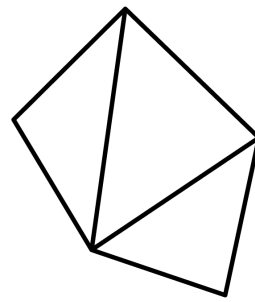
Abbildung 2.7.: Transformation mittels Halfedge Collapse

Vertex Removal

Hierbei wird ein Vertex entfernt und das Resultat neu trianguliert. Triangulation ist ein Verfahren um ein Polygon in *Triangles* aufzuteilen. In Abbildung 2.8a wird der zentrale Vertex entfernt. Anschliessend werden alle *Triangles* an diesem Punkt entfernt. Das dabei entstehende Loch wird neu mit *Triangles* gefüllt. Der neu triangulierte Polygon ist in Abbildung 2.8b ersichtlich.



(a) Modell Ausgangslage



(b) Vertex Removal

Abbildung 2.8.: Transformation mittels Vertex Removal

2.5. Graphics Pipeline

Die Graphics Pipeline ist zuständig um eine definierte Szene auf einem Ausgabegerät zu visualisieren. Die verschiedenen Schritte werden im folgenden Abschnitt kurz erläutert. Es wird dabei ein simples 3D-Modell auf einen normalen 2D Display visualisiert. Die verschiedenen Schritte sind in Abbildung 2.9 aufgezeigt.



Abbildung 2.9.: Schritte einer Graphics Pipeline

Applikation

Im ersten Schritt wird die Szenerie aufbereitet, Modelle in den Arbeitsspeicher geladen, Positionen und Rotationen der Modelle definiert. Diese Schritte werden auf der CPU durchgeführt.

Geometrie

Anschliessend werden die Definitionen an die Geometrie Pipeline geliefert. Diese ist zuständig um die Modelle auszurichten, die Beleuchtung zu bestimmen und die 2D Projektion vorzunehmen.

Rasterization

Im letzten Schritt werden kontinuierliche Objekte zu diskreten Fragmenten verarbeitet. Der Anwender hat hierbei in vielen Engines die Möglichkeit, den Prozess mithilfe eines Fragmentshaders ³ anzupassen.

Bildschirm

Am Schluss wird das vom Rasterization Prozess generierte Bild auf dem Bildschirm dargestellt.

2.6. Performanzoptimierung

Visualisierungen können zu komplex werden, um jederzeit performant und interaktiv gerendert zu werden. Insbesondere wenn viele Objekte gleichzeitig sichtbar sind, lohnt es sich Performanzoptimierungen durchzuführen. Im Idealfall geschieht dies jedoch, ohne dass der Anwender dies bemerkt.

In diesem Abschnitt werden mögliche Ansätze erklärt, welche helfen sollen, die Render-Performanz zu erhöhen. Diese Arbeit konzentriert sich jedoch auf den Ansatz von Level of Detail; die anderen Ansätze werden nur kurz erläutert. Detailliert auf Level of Detail wird in Abschnitt 2.7 eingegangen.

Frustum culling

Polygone, welche nicht im Kamera Frustum enthalten sind, werden bei dieser Methode nicht weiter prozessiert. Dies reduziert die Anzahl Polygone drastisch.

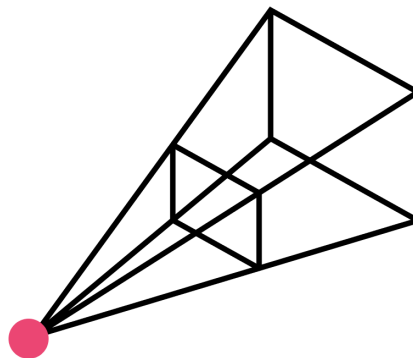


Abbildung 2.10.: Kamera Frustum

³Generiert Farbdefinition für ein einzelnes Pixel mit zugehörigem Tiefenwert

Ein Frustum ist eine geometrische Form, die einer Pyramide ähnelt, welcher die Spitze abgeschnitten wurde. Das Kamera-Frustum bezeichnet den Raum, welcher von der Kamera aufgezeichnet wird. Dieser Bereich startet eine gewisse Distanz von der Kamera entfernt und endet weit hinten wieder, da nicht bis ins unendliche Objekte sichtbar sind. Wie in Abbildung 2.10 zusehen ist, beginnt das Kamera-Frustum in der Nähe der Kamera und dehnt sich ein wenig aus bis es gegen Ende des Spektrums beschränkt wird.

Occlusion culling

Polygone bzw. Objekte, welche komplett von anderen Objekten überdeckt werden, werden bei dieser Variante nicht prozessiert.

Backface culling

Bei dieser Methode wird berechnet welche Polygone zur Kamera orientiert sind. Alle Polygone, welche in die andere Richtung zeigen werden nicht gezeichnet. Dies ist nicht immer gewünscht, für die meisten Anwendungen ist diese Optimierung jedoch aktiviert. Als Grundlage für die Berechnung werden die *Normals* der *Vertices* berücksichtigt.

Parallel rendering

Auch bekannt unter Distributed Rendering ist der Einsatz von Techniken aus dem Parallel Programming in Visualisierungsanwendungen.

Image-based rendering

In gewissen Fällen kann das Modellieren übersprungen werden. So kann anhand von Bildmaterial eine 3D-Illusion erzeugt werden.

2.7. Einführung Level of Detail

Als Level Of Detail (LOD) werden die verschiedenen Detailstufen bei der virtuellen Darstellung bezeichnet. Dies wird verwendet, um die Geschwindigkeit von Anwendungen zu steigern, indem Objekte im Nahbereich detailliert angezeigt werden; wohingegen Elemente im Fernbereich deutlich vereinfacht dargestellt werden.

Ein LOD Algorithmus hat das Ziel für ein gegebenes Modell eine vereinfachte Darstellung zu finden, welche das Original ausreichend annähert. Um diese Approximationen zu generieren, kann eine Vielzahl von Algorithmen verwendet werden. Es gibt verschiedene Ansätze zur Generierung von LODs, welche in Unterabschnitt 2.7.2 im Detail erläutert werden.

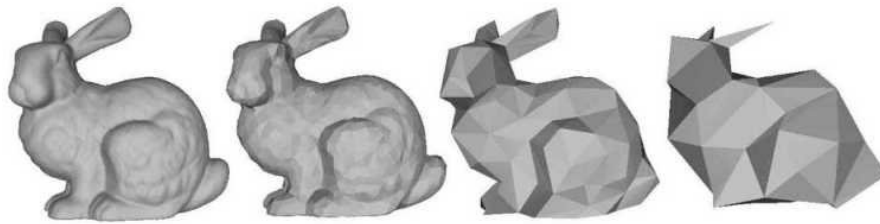


Abbildung 2.11.: Level Of Detail Visualisierung vier Hasen

Wie in der Abbildung 2.11 zu erkennen ist, wird von links nach rechts der Detailgrad und somit die Komplexität des Objektes reduziert. Sind es im Bild ganz links noch 69'451 Polygone, wird es bereits im ersten Schritt auf 2'502 Polygone reduziert. Dies ist eine enorme Reduktion von ca. 96.5%. Im dritten Schritt wird die Anzahl Polygone wiederum um ca. 90% auf 251 reduziert. Schlussendlich hat das letzte Objekt noch 76 Polygone was knapp 0.1% der ursprünglichen Anzahl entspricht.

2.7.1. Ansätze für LOD Artefakte

Es gibt verschiedene Ansätze, 3D-Modelle mittels LOD zu vereinfachen. In diesem Abschnitt werden einige davon detaillierter erläutert so wie ihre Vor- und Nachteile aufgezeigt.

Diskrete LOD (DLOD)

Bei diskreten LOD werden für ein detailliertes Modell mehrere weniger detaillierte Modelle erstellt. Abhängig von der Distanz zum Betrachter wird das optimale Modell gewählt. Es sind hierfür nur minimale Anpassungen am Scene Graph ⁴ notwendig da ausschliesslich ganze Modelle ausgetauscht werden. Ein Nachteil sind jedoch merkbare harte Grenzen. Der Benutzer stellt beim umhergehen in der Szene fest, wenn das Modell mit einer einfacheren Version ausgetauscht wird. Es ist zudem häufig nicht möglich grosse Modelle sinnvoll zu vereinfachen. So ist der Ansatz für Terrain ungeeignet. Ähnlicherweise ist es auch nicht möglich viele sehr kleine Modelle zu kombinieren da jedes Modell unabhängig ist.

Kontinuierliche LOD (CLOD)

Im Gegensatz zu DLOD wird bei CLOD vereinfachende Veränderungen an einem Modell gespeichert.

Der Hauptunterschied zu DLOD besteht in den weichen Grenzen, so ist es signifikant weniger auffällig da zwischen den verschiedenen Auflösungen interpoliert werden kann. Der Hauptnachteil ist jedoch, dass diese Interpolation insbesondere Auswirkungen auf die Laufzeit Performanz der Applikation hat. Das Problem des Clusterings ist auch für diese Variante nicht gelöst.

⁴Objektorientierte Datenstruktur zur Beschreibung von 2D- oder 3D-Szenarien

Hierarchische LOD (HLOD)

Bei HLOD werden mehrere Objekte in einen Cluster gruppiert. Diese Methode erlaubt es somit zum Beispiel Terrains optimal abzubilden. So werden Ausschnitte, welche nahe beim Benutzer sind mit hoher Auflösung dargestellt während weiter entfernte Teile der Umgebung mit weniger Details ausgestattet sind. Zudem ist es möglich viele kleinere Objekte bei entsprechenden Distanzen zu kombinieren. Der Hauptnachteil liegt darin, dass hierfür signifikante Anpassungen am Scene Graph notwendig sind und für den Entwickler ein spürbarer Unterschied entstehen kann.

2.7.2. Vergleich Algorithmen

Diese Art von Problem ist in der Literatur unter anderem als *Surface Simplification*, polygonale Simplifizierung, geometrische Simplifizierung oder Mesh Reduzierung bekannt. Grundsätzlich kann man die verschiedenen Algorithmen in Kategorien einteilen. Im folgenden Abschnitt wird auf die Grundidee der verschiedenen Kategorien eingegangen.

Vertex Dezimierung

Algorithmen dieser Kategorie entfernen jeweils einen Vertex und triangulieren das entstehende Loch neu. Eine weitere Möglichkeit ist das *Vertex Clustering*, hierbei wird ein Modell in verschiedene Gitterzellen aufgeteilt. Die Punkte innerhalb einer Gitterzelle werden dann in einen einzigen Punkt zusammengelegt. Die visuelle Annäherung dieser Algorithmen bei drastischen Vereinfachungen stellt jedoch häufig ein Problem dar für polygonale Modelle. Sogenannte Point Clouds ⁵, wie sie bei 3D-Scannern eingesetzt werden, können mit solchen Algorithmen schnell vereinfacht werden.

Edge Collapse

Hierbei werden iterativ *Edges* entfernt und die beiden *Vertices* zu einem Punkt zusammengelegt. Die verschiedenen Algorithmen unterscheiden sich primär in der Selektion der *Edges*. Grundsätzlich wird hierfür eine Heuristik verwendet um den Fehler quantifizieren zu können. Anschliessend werden iterativ die *Edges* entfernt, welche zu einem minimalen Fehler führen. Beim Zusammenführen von *Edges* sind abhängig vom gewählten LOD System unterschiedliche Strategien gefordert. Bei diskreten LOD Systemen soll nach dem Zusammenführen ein optimaler neuer Punkt gefunden werden, bei kontinuierlichen LOD Systemen wird häufig einer der beiden Punkte gewählt um Speicherplatz einsparen zu können.

Face Collapse

Bei dieser Variante werden jeweils Flächen entfernt und die umliegenden Flächen zusammengelegt.

⁵Vertices im dreidimensionalen Raum ohne zugehörige *Triangle*

3. Vorgehen

3.1. Nutzen LOD

Um den Nutzen von LOD quantifizieren zu können, wird in einer ersten Phase ein Benchmark aufgestellt. Ziel ist es, das Laufzeitverhalten unter Einsatz eines optimierten Modells zu analysieren und somit den maximal möglichen Einfluss von LOD auf die Leistung klassifizieren zu können. Angestrebt werden 60 FPS¹. Wenn die unoptimierte Variante diesen Wert erreicht oder gar übersteigt, ist eine Optimierung nicht notwendig. Erreicht die optimierte Variante diesen Wert nicht, ist die Szenerie nach wie vor zu komplex. Ziel der Arbeit ist es, einen Algorithmus zu entwickeln, der für Szenarien unter 60 FPS diese erhöht und im Idealfall auf ebendiesen Wert hebt.

3.1.1. Definition Benchmark

In diesem Abschnitt werden die verschiedenen erwägten Optionen erläutert und die Abgrenzungen aufgezeigt.

Browser Umgebung

Um den Umfang des Benchmarks überschaubar zu halten, wurde ausschliesslich ein Benchmark für Google Chrome entwickelt. Google Chrome basiert auf Chromium², dieselbe Engine, welche auch Microsoft Edge oder Opera verwenden. Einen Benchmark basierend auf Google Chrome liefert somit auch Indizien für diese beiden Browser, auch wenn gewisse Abweichungen möglich sind. Neben dem Marktführer Chrome sind Mozilla Firefox oder Safari von Apple ebenfalls Optionen. Jedoch wurde primär aufgrund des Marktanteils von total rund 70% [6] zugunsten von Google Chrome entschieden. Die getroffenen Aussagen bezüglich Laufzeitverhalten behalten ihre Gültigkeit auch für andere Browser.

Automation

Um die Tests durchzuführen, wird ein Testautomationstool benötigt; unter anderem der Einsatz von Selenium wurde in Erwägung gezogen. Der Vorteil von Selenium ist insbesondere, dass der Benchmark für weitere Browser ausgeweitet werden könnte. Da jedoch das Analysieren der GPU Daten stark vom System abhängig ist und dafür zusätzliche

¹Frames per Second. Bilder pro Sekunde. Gängige Art, die Performanz der Rendering Engine zu messen

²Open Source Browser-Projekt. Grundlage von Browsern wie Google Chrome, Edge oder Opera

Komplexität notwendig wäre, wird in diesem Benchmark die im Google Chrome integrierten *Chrome DevTools* eingesetzt. Selenium bietet zurzeit eine suboptimale Integration für das *Chrome DevTools Protocol*. Eine unabhängige Arbeit definierte 2017 einen vergleichbaren Benchmark um die Performanz von Three.js und Babylon.js in einem spezifischen Anwendungsgebiet zu vergleichen. Der Benchmark misst die Systemdaten unabhängig von der Browser Instanz. [7] Um mögliche Diskrepanzen zwischen Systemen möglichst gering zu halten wurde jedoch entschieden auf die bewährte Lösung von Google Chrome zu setzen. Puppeteer³, eine weitere Option für die Automation, ist eine Bibliothek, die eine vereinfachte Schnittstelle zu einer Headless Chrome Instanz bietet. Sie wird zudem direkt von Google entwickelt und bietet somit eine stabile Grundlage zur Kommunikation mit den *Chrome DevTools*.

Profiling

Die *Chrome DevTools* erlauben es, ein detailliertes Laufzeitprofil einer Applikation anzulegen. Im Profil befinden sich Informationen zu CPU- und GPU-Auslastung aber auch generelle Informationen bzgl. der Rendering Engine werden gesammelt. Die Analyse dieser Daten ermöglicht es, eine Aussage zum Laufzeitverhalten einer Applikation zu tätigen.

Testaufbau

Derselbe Testablauf wird sowohl für die optimierte als auch für die unoptimierte Version verwendet. Bei einem Testablauf werden folgende Schritte durchlaufen:

1. Öffne Applikation in *Headless Chrome* Instanz.
2. Warte bis Seite geladen ist.
3. Starte *Profiling*.
4. Warte n Sekunden.
5. Stoppe *Profiling*.
6. Werte Daten aus.

Der Test erfasst kontinuierlich die *FPS*. Zudem wird aus dem *Profiling* eine Kennzahl bzgl. der GPU Nutzung berechnet.

Aufbau Testapplikation

Die Testapplikation stellt eine komplexe Szenerie dar. Der Betrachter fliegt während dem Ablauf kontinuierlich über die Szenerie. Dies stellt die optimalen Bedingungen für den Einsatz von LOD Artefakten dar.

³Node.js Library, die eine API anbietet zum Steuern von Chromium oder Chrome über das Chrome DevTools Protocol

Testumgebung

Um während den Tests möglichst faire Bedingungen zu gewährleisten wird die Maschine zuvor wie bei anderen Benchmarks vorbereitet. Ziel ist es, Seiteneffekte zu minimieren. Für diesen Benchmark wurden deshalb die Instruktionen von *Tracer Bench* zur Behebung von Rauschen befolgt. [8]

Analyse der Daten

Um eine zuverlässige Aussage treffen zu können, wird der Vorgang mehrfach wiederholt. Für jeden Durchlauf wird der Median der *FPS* Daten berechnet. Anschliessend wird die Varianz der *FPS* für die unoptimierten respektive optimierten Werte berechnet. Die Varianz dient als Kennzahl, um eine Signifikanz der Daten nachweisen zu können.

Konfidenzintervall Die Signifikanz wird mithilfe eines statistischen Konfidenzintervalls nachgewiesen. Hierfür wird der Durchschnitt aller Mediane verwendet, zusätzlich wird ein Konfidenzintervall von 95% gewählt. Somit wird gewährleistet, dass das Resultat des Benchmarks verlässlich ist und für einen Vergleich verwendet werden kann.

3.2. Vergleich LOD Systeme

Die verschiedenen LOD Systeme sowie ihre Vor- und Nachteile werden in Unterabschnitt 2.7.1 erläutert. In diesem Schritt wird erläutert, welche Art LOD System in dieser Arbeit eingesetzt werden soll.

Die verschiedenen Ansätze unterscheiden sich in einigen Kriterien. Für den Einsatz in einer Webanwendungen ist die Gewichtung der Kriterien anders als für Desktopanwendungen, Spiele oder dergleichen. So hat die Dateigrösse der Modelle bei Webanwendungen einen grösseren Einfluss auf die gefühlte Performanz beim Endanwender. Insofern ist die Auswirkung der gewählten Strategie auf die Downloadgrösse ein wichtiges Kriterium. Für diskrete LOD muss für jedes Level ein separates Modell geladen werden. Dies erhöht die Downloadgrösse merkbar. Bei kontinuierlichen LOD Systemen muss das Modell die Unterschiede zwischen verschiedenen Levels persistieren. Mit einem für kontinuierliche LOD ausgelegten Algorithmus kann der benötigte Speicher reduziert werden, indem beim *Edge Collapse* nicht auf einen neuen optimalen *Vertex* reduziert wird, sondern von den beiden bestehenden *Vertices* der passendere gewählt wird. Trotzdem entsteht für jeden *Edge Collapse* ein zusätzlicher Speicheraufwand. Insbesondere unter Berücksichtigung der Tatsache, dass Laufzeitperformanz essentiell für Webanwendungen ist, wurde entschieden nicht auf kontinuierliche LOD Systeme zu setzen. Zudem wurde entschieden kein hierarchisches LOD System zu entwickeln, um die Integration des Tools für die Entwickler einer Webanwendung möglichst einfach zu gestalten. Das diskrete LOD System ist für Grafiker und Entwickler die einfachste zu verwenden und deckt somit auch die breiteste Benutzerbasis und die meisten Anwendungsfälle ab. Aufgrund dessen, wird im Rahmen dieser Arbeit ein Algorithmus dieses Systems entwickelt und in den Entwicklungsprozess integriert. Eine Übersicht über die verschiedenen Kriterien kann in Tabelle 3.1 gefunden werden.

Technische Auswirkungen	DLOD	CLOD	HLOD
Auswirkung auf Downloadgrösse	mittel	mittel	mittel
Auswirkung auf Laufzeitverhalten	klein	mittel	mittel
Auswirkung auf <i>Scene Graph</i>	lokal	lokal	global
"Visual Pop" vermeidbar	nein	ja	ja
Integration in bestehende Arbeitsabläufe	einfach	einfach	umständlich
Möglichkeiten			
Drastische Reduktion von Polygonen	ja	ja	ja
Clustering möglich	nein	nein	ja

Tabelle 3.1.: Übersicht Merkmale der verschiedenen LOD Systeme

3.3. Surface Simplification Algorithmus

Die verschiedenen Klassen von Algorithmen werden in Unterabschnitt 2.7.2 kurz erläutert.

Verschiedene Algorithmen wurden für die Implementierung in Betracht gezogen. Die wohl weitverbreitetste Art ist der von Garland und Heckbert 1997 definierte *Surface Simplification using Quadric Error Metrics* welcher auf *Edge Collapses* basiert. [9]

Der Algorithmus kann wie folgt zusammengefasst werden:

1. Fehler Metrik für alle Vertices berechnen.
2. Optionen für *Edge Collapse* markieren.
3. Optionen mit geringstem Fehler optimieren.

3.3.1. Fehler Metrik

Das Ziel der Fehler Metrik ist es eine Heuristik zu definieren, welche den geometrischen Unterschied zwischen dem vereinfachten und dem originalen Modell beschreibt. Mithilfe der Fehler Metrik können anschliessend die Transformationen am Modell gemäss den kleinsten Kosten optimiert werden. Der Algorithmus funktioniert grundsätzlich unabhängig von der gewählten Fehler Metrik. So könnten verschiedene Metriken eingesetzt werden. Die hier gewählte Metrik definiert den Fehler mithilfe einer symmetrischen 4×4 Matrix. Die Matrix repräsentiert die Distanz eines *Vertices* zu einer zugehörigen *Face*. Für die Kombination mehrerer *Faces* können die Matrizen addiert werden.

Beim Entfernen einer *Edge* werden die Fehler Metriken der zugehörigen *Vertices* addiert.

3.4. Pipeline Integration

Die Lösung soll in eine wiederverwendbare und konfigurierbare Pipeline integriert werden. In der Webentwicklung ist es üblich, über CLI⁴ Arbeitsschritte zu automatisieren.

⁴Command Line Interface – Kommandozeile

Es soll demnach ein *CLI* erstellt und auf NPM ⁵ öffentlich zugänglich gemacht werden, welches nahtlos in den Entwicklungsablauf integriert werden kann. Über eine Konfiguration soll angegeben werden können, wo nach den *.glTF* Dateien gesucht, wie die Output-Dateien benannt und in welchem Modus das Tool gestartet werden soll. Ebenso werden Algorithmus-Einstellungen in dieser Konfiguration gesetzt werden können. Es soll einfach zu bedienen, gut dokumentiert und gut gewählte Standardeinstellungen haben. Zudem soll das Tool im einmal Modus laufen können oder kontinuierlich sich ändernde Dateien neu optimieren.

3.5. Automatische Generierung von Detail Levels

Damit für den Endbenutzer die Konfiguration übersichtlich bleibt, soll der Einsatz von zum Beispiel heuristischen Methoden hilfreiche Basiskonfigurationen liefern. Zum einen geht es hier um das Festlegen der Thresholds, aber auch das Definieren des Dezimierungs- / Vereinfachungsfaktors spielt eine wichtige Rolle.

3.6. Levelwahl während Laufzeit

Die optimale Levelwahl kann stark von der Laufzeitumgebung abhängig sein. So sollte zum Beispiel auf mobilen Geräten früher ein einfacheres Modell gezeigt werden als auf leistungsstarken Geräten. Deshalb eignet sich eine Wahl der Levels zur Laufzeit, um ein optimales Erlebnis auf allen Geräten zu ermöglichen. Nebst dem *CLI* soll folglich auch ein Javascript-Modul bereitgestellt werden, welches das Laden von glTF Dateien übernimmt und basierend auf der erkannten Performanz das beste Level of Detail einspielen.

⁵Node Package Manager

4. Resultate

Das Resultat besteht wie beschrieben aus verschiedenen Artefakten. In diesem Abschnitt wird konkret auf die Resultate der verschiedenen Schritte eingegangen und der Zusammenhang hergestellt.

4.1. Benchmark

Der Benchmark dient als Mittel um den Nutzen von LOD in Webanwendungen quantifizieren zu können. Die dafür zur Hand genommene Masseinheit ist FPS.

4.2. LOD Pipeline

Der Kern der Arbeit stellt die LOD Pipeline, auch *lode* genannt, dar. Dieses Tool liefert eine für das Web optimierte Möglichkeit LOD Artefakte für eine Vielzahl von Anwendungsgebieten zu generieren.

5. Diskussion und Ausblick

Das entwickelte Tool, *lode*, liefert einen ersten Schritt in die Richtung einer LOD Pipeline, welche für das Web optimiert wurde. Das Tool erreicht das Ziel für eine Vielzahl von Anwendungsgebieten einsetzbar zu sein. Wichtig ist es jedoch festzuhalten, dass es nicht für alle Anwendungen sinnvoll ist. Zum einen soll die Pipeline weiter entwickelt werden und in ihren Kernkompetenzen gestärkt werden, zum anderen gibt es jedoch auch spezifische Gebiete wie zum Beispiel Terrain für welches *lode* ungeeignet ist.

6. Verzeichnisse

Glossar

Chromium Open Source Browser-Projekt. Grundlage von Browsern wie Google Chrome, Edge oder Opera. 19, 20, 27

CLI Command Line Interface – Kommandozeile. 22, 23

FPS Frames per Second. Bilder pro Sekunde. Gängige Art, die Performanz der Rendering Engine zu messen. 19–21, 24

Fragments shaders Generiert Farbdefinition für ein einzelnes Pixel mit zugehörigem Tiefenwert. 15

Game Engines Framework, das für den Spielverlauf und dessen Darstellung verantwortlich ist. 6

NPM Node Package Manager. 23

OpenGL Spezifikation einer Programmierschnittstelle zur Entwicklung von 2D und 3D-Grafikanwendungen. 7

Point Cloud Vertices im dreidimensionalen Raum ohne zugehörige *Triangle*. 18

Puppeteer Node.js Library, die eine API anbietet zum Steuern von Chromium oder Chrome über das Chrome DevTools Protocol. 20

Rendering Engine Teilprogramm, das zuständig für die Darstellung von Grafiken ist. 6, 19, 20, 27

Scene Graph Objektorientierte Datenstruktur zur Beschreibung von 2D- oder 3D-Szenarien. 17

Texture Mapping Verfahren, mittels 2D-Bildern 3D-Objekte zu gestalten. 9

Literaturverzeichnis

- [1] D. Jackson und J. Gilbert. (2021) *WebGL Specification*. [Online]. URL: <https://www.khronos.org/registry/webgl/specs/latest/1.0/> [Stand: 26.03.2021].
- [2] GPU for the Web Community Group. (2021) *GPU for the Web Community Group Charter*. [Online]. URL: <https://gpuweb.github.io/admin/cg-charter.html> [Stand: 26.03.2021].
- [3] M. Segal, K. Akeley, C. Frazier, J. Leech und P. Brown. (2008) *The OpenGL® Graphics System: A Specification*. [Online]. URL: <https://www.khronos.org/registry/OpenGL/specs/gl/glspec30.pdf> [Stand: 26.03.2021].
- [4] unknown. (2007) *Object Files*. [Online]. URL: <http://www.martinreddy.net/gfx/3d/OBJ.spec> [Stand: 26.03.2021].
- [5] P. Cozzi und T. Parisi. (2016) *glTF 1.0 Specification*. [Online]. URL: <https://github.com/KhronosGroup/glTF/tree/master/specification/1.0> [Stand: 26.03.2021].
- [6] StatCounter. (2021) *Browser Market Share Worldwide / StatCounter Global Stats*. [Online]. URL: <https://gs.statcounter.com/browser-market-share#monthly-202003-202103> [Stand: 14.04.2021].
- [7] O. Nordquist und A. Karlsson. (2017) *BabylonJS and Three.js - Comparing performance when it comes to rendering Voronoi height maps in 3D*. [Online]. URL: <http://www.diva-portal.org/smash/get/diva2:1228221/FULLTEXT01.pdf> [Stand: 17.04.2021].
- [8] M. Lynch. (2020) *TracerBench: Noise Mitigation Techniques*. [Online]. URL: https://github.com/TracerBench/tracerbench/blob/master/NOISE_MITIGATION.md [Stand: 17.04.2021].
- [9] M. Garland und P. S. Heckbert. (1997) *Surface Simplification Using Quadric Error Metrics*. [Online]. URL: <https://www.cs.cmu.edu/~./garland/Papers/quadrics.pdf> [Stand: 17.04.2021].

Abbildungsverzeichnis

2.1. Vereinfachung von Polygone als Sammlung von <i>Triangles</i>	8
2.2. Convertierungspipeline vor glTF [5]	10
2.3. Convertierungspipeline mit glTF [5]	10
2.4. glTF Datenstruktur [5]	11
2.5. Modell Basis	12
2.6. Transformation mittels Edge Collapse	12
2.7. Transformation mittels Halfedge Collapse	13
2.8. Transformation mittels Vertex Removal	14
2.9. Schritte einer Graphics Pipeline	14
2.10. Kamera Frustum	15
2.11. Level Of Detail Visualisierung vier Hasen	17

Tabellenverzeichnis

3.1. Übersicht Merkmale der verschiedenen LOD Systeme	22
---	----

A. Anhang

A.1. Aufgabenstellung

Aufgabenstellung

A.2. Anhang 2

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.