

**Bachelorarbeit**

**im Studiengang Medientechnik**

Eingereicht von: Enis Inep

Matrikelnummer: 7022248

Mail: [enis.inep@stud.hs-emden-leer.de](mailto:enis.inep@stud.hs-emden-leer.de)

Erstprüfer:

Prof. Dr. rer. nat. Thies Pfeiffer

Zweitprüfer:

Jannik Franssen, M.Eng.

Hamburg, 09 2024

Inhaltsverzeichnis

[1 Einleitung 1](#_Toc210174366)

[1.1 Motivation und Relevanz des Themas 1](#_Toc210174367)

[1.2 Problemstellung: Tiefenwahrnemung bei Panoramabildern auf 2D Monitoren 1](#_Toc210174368)

[1.3 Zielsetzung und Forschungsfragen 1](#_Toc210174369)

[1.4 Aufbau der Arbeit 1](#_Toc210174370)

[2 Theoretische Grundlagen 4](#_Toc210174371)

[2.1 Wahrnehmungspsychologische Grundlagen (Hergovich, 2022a) 4](#_Toc210174372)

[2.1.1 Parallax Effect und Tiefenhinweise (monokular vs. binokular)(Wölfel, 2023) 4](#_Toc210174373)

[2.1.2 Grenzen der Tiefenwahrnehmung auf 2D-Displays 5](#_Toc210174374)

[2.2 360°-Panoramen als Darstellungsformat 5](#_Toc210174375)

[2.2.1 Equirectangular Projection und Verzerrungen 5](#_Toc210174376)

[2.2.2 Rolle von Depth Maps (Funktionsweise, Nutzen, Limitationen) 6](#_Toc210174377)

[2.3 Überblick bestehender Ansätze 7](#_Toc210174378)

[2.4 Stand der Forschung im Bereich Tiefenwahrnemung 7](#_Toc210174379)

[3 Methodik 8](#_Toc210174380)

[3.1 Forschungsdesign: Analyse und Bewertung der Szenarien 8](#_Toc210174381)

[3.2 Definition der Szenarien 8](#_Toc210174382)

[3.2.1 Szenario 1: Baseline-3D-Rendering mit gezielten Tiefenherausforderungen 8](#_Toc210174383)

[3.2.2 Szenario 2: Depth Map + 360°-Panorama (*Babylon.js*, Artefakte, Optimierungsstrategien) 8](#_Toc210174384)

[3.2.3 Szenario 3: Low-Poly-Rekonstruktion aus Panorama + ggf. Depth Map (Tools wie zForm, HoHoNet) 9](#_Toc210174385)

[3.3 Technische Umsetzung 9](#_Toc210174386)

[3.4 Evaluationsdesig 9](#_Toc210174387)

[3.4.1 Stimulusmaterial 9](#_Toc210174388)

[3.4.2 Auswertungsmethode 9](#_Toc210174389)

[4 Ergebnisse 10](#_Toc210174390)

[4.1 Technische Ergebnisse der Szenario-Umsetzungen 10](#_Toc210174391)

[4.2 Beobachtete Optimierungseffekte (z. B. weniger Artefakte, visuelle Glätte) 10](#_Toc210174392)

[4.3 Ergebnisse der Nutzerbefragung (Darstellung in Tabellen, Diagrammen) 10](#_Toc210174393)

[4.4 Zusammenfassung der Ergebnisse in Bezug auf Forschungsfrage 10](#_Toc210174394)

[5 Diskussion 11](#_Toc210174395)

[5.1.1 Interpretation der Ergebnisse im Kontext der Forschungsfrage 11](#_Toc210174396)

[5.1.2 Vergleich mit Literatur(?) 11](#_Toc210174397)

[5.1.3 Bewertung der Verfahren: Stärken, Grenzen, Anwendungspotenzial 11](#_Toc210174398)

[6 Fazit und Ausblick 12](#_Toc210174399)

[6.1 Beantwortung der Forschungsfrage 12](#_Toc210174400)

[6.2 Beitrag der Arbeit 12](#_Toc210174401)

[6.3 Offene Fragen und Ausblick 12](#_Toc210174402)

[7 Anhang XIII](#_Toc210174403)

# Einleitung

Im Rahmen meines Praktikums bei einem VR/XR-Start-up in Hamburg habe ich ein ausgeprägtes Interesse an der Tätigkeit als 3D Artist entwickelt. Nach Abschluss des Praktikums habe ich mich am Mixed-Reality-Lab der Hochschule Emden/Leer nach einem geeigneten Thema für eine Bachelorarbeit erkundigt. Durch meinen Betreuer, Herrn Dr. Prof. Theis Pfeiffer, wurde ich darüber informiert, dass das Mixed-Reality-Lab regelmäßig mit 360°-Panoramen im Bereich der virtuellen Realität arbeitet. Diese Bilder enthalten keine Tiefendaten und wirken daher zweidimensional. Mithilfe einer Depth Map ist es möglich, ein Tiefeneindruck zu erzeugen, der dem Panorama mehr Tiefe verleiht. Allerdings weist diese Methode auch einige Schwächen auf. Das Verfahren könnte optimiert werden, um diesen Tiefeneindruck zu erzeugen. Im Rahmen der Analyse werden drei Herangehensweisen erörtert und umgesetzt. Darüber hinaus interessieren mich auch Bilder und Computergrafik im allgemeinen und daher ist das Thema für mich von großem Interesse.

## Motivation und Relevanz des Themas

Wir sind motiviert, weil wir motiviert sind

## Problemstellung: Tiefenwahrnemung bei Panoramabildern auf 2D Monitoren

## Zielsetzung und Forschungsfragen

Das zentrale Ziel dieser Arbeit ist der Versuch, mithilfe von drei unterschiedlichen Darstellungsverfahren ein visuell überzeugendes Ergebnis für die Wahrnehmung von Tiefe auf einem 2D-Monitor zu erzielen. Die zugrunde liegende Forschungsfrage lautet daher: Wie lassen sich drei Verfahren zur Darstellung von Panoramabildern auf 2D-Monitoren technisch korrekt und visuell optimiert umsetzen, um eine möglichst überzeugende Tiefenwahrnehmung zu erzielen, und welche Grenzen zeigen sie dabei. Somit könnte ein einstieg in die Weiterführende Optimierung vom Tiefeneindruck, für Panoramabilder geschafft werden.

## Aufbau der Arbeit

Die drei Herangehensweisen werden in einer 3D-Umgebung erstellt und jedes Szenario wird individuell optimiert.

**Szenario 1: S1**

Das Basisszenario S1 ist ein 3D-Rendering, in dem Objekte bewusst so angeordnet sind, dass Grenzen der Tiefenwahrnehmung ausgereizt werden. Die Herausforderung besteht darin, dass Nutzer:innen bei der Betrachtung Schwierigkeiten haben können, Tiefe, Abstände und Relationen korrekt einzuschätzen. Damit dient S1 als Testszene, um systematisch zu prüfen, welche Probleme bei der Wahrnehmung von Tiefe auftreten.

Aufbauend auf S1 werden weitere Szenarien erzeugt. Diese stellen verschiedene technische Lösungsansätze dar, um die Wahrnehmung von Tiefe zu verbessern. Ein Highpoly-Rendering wie in S1 zeigt zwar hohe Detailtreue, bringt aber Herausforderungen hinsichtlich Performance und Dateigröße mit sich, insbesondere in einer Web-Umgebung.

**Szenario 2: S2**

Szenario 2 S2 erweitert die Basisszene S1 durch die Fusion einer Depth Map mit dem 360°-Panorama. Mithilfe von Babylon.js wird auf einer unterteilten Sphere ein Tiefeneindruck erzeugt. Jedes Face erhält den Tiefenwert seines Mittelpunktes, wodurch die Szene räumlich dargestellt wird.

Bei der Analyse von S2 treten jedoch neue Herausforderungen auf. Insbesondere an Spitzen, Ecken und Außenkanten entstehen Verzerrungen, wenn sich die Nutzer:innen im Raum bewegen oder drehen. Dies liegt daran, dass die Kanten der Faces nach hinten abknicken und in der Mitte nadelartige Spitzen entstehen, die den Gesamteindruck der Szene beeinträchtigen.

Ein Bild, das Handschrift, Text, Entwurf, Diagramm enthält.

KI-generierte Inhalte können fehlerhaft sein.Ein Bild, das Handschrift, Entwurf, Schrift, Zeichnung enthält.

KI-generierte Inhalte können fehlerhaft sein.

Abbildung 1 Grobe Skizze Szenario 2

Szenario 3 S3 verfolgt den Ansatz, ein Lowpoly-Mesh aus der Basisszene S1 zu erzeugen, das mit gebackenen Texturen versehen werden kann. Ziel ist es, ein visuell zufriedenstellendes Ergebnis zu erreichen, ohne ein neues Mesh manuell zu modellieren. Im Unterschied zu S2 soll das Mesh hier nicht zwingend als durchgehendes, geschlossenes Modell umgesetzt werden. Vielmehr bietet sich die Möglichkeit, Objekte im Raum vom Hintergrund zu trennen, sodass unnötige Flächen insbesondere hinter dem eigentlichen Mesh eingespart werden können.

Da die konkrete Herangehensweise für S3 noch nicht endgültig festgelegt ist, besteht eine Herausforderung darin, eine geeignete Methode zur Lowpoly-Erzeugung zu identifizieren. Dabei müssen einerseits visuelle Qualität und Tiefeneindruck gewahrt bleiben, andererseits dürfen Performance und Dateigröße nicht aus dem Ruder laufen. Die Trennung von Objekten und Hintergrund erfordert zudem eine saubere Segmentierung, um Darstellungsfehler zu vermeiden. Für eine fundierte Bewertung wird es sinnvoll sein, die Ergebnisse online oder in Form von Videodemonstrationen zu präsentieren, damit der Parallax-Effekt während der Beurteilung berücksichtigt werden kann.

Ein Bild, das Handschrift, Text, Schrift, Tinte enthält.

KI-generierte Inhalte können fehlerhaft sein.

Abbildung 2 Grobe Skizze Szenario 3

# Theoretische Grundlagen

Das Thema der Untersuchung lautet: "Untersuchung dreier Verfahren zur Darstellung von Panoramabildern mit visuell optimierter Tiefenwahrnehmung und ihren Grenzen". Für die Untersuchung ist ein gewisses Maß an Verständnis der Wahrnehmungspsychologie erforderlich. Es stellt sich die Frage, wie Tiefe auf einem 2D-Monitor wahrgenommen wird, wann genau Tiefe erkannt wird, was der Parallaxe-Effekt ist und wozu er wichtig ist. Zudem wird erörtert, wozu DepthMaps dienen und welche Funktionen sie bieten.  
Im Folgenden erfolgt eine Definition des Begriffs *360°-Panorama*. Ein 360°-Panorama liegt in equirektangularer Projektion vor. Dabei wird die sphärische Umgebung in eine zweidimensionale, rechteckige Fläche mit einem Seitenverhältnis von 2:1 transformiert. Dieses Format kann in 3D-Anwendungen wieder auf eine Kugelgeometrie projiziert werden, wodurch eine immersive Rundumsicht entsteht. In der vorliegenden Arbeit wird ein direkt anschließendes, umlaufendes Bild als 360°-Panorama bezeichnet.

## Wahrnehmungspsychologische Grundlagen (Hergovich, 2022a)

Unnötiges Kapitel

### Parallax Effect und Tiefenhinweise (monokular vs. binokular)(Wölfel, 2023)

Wie bereits in den „Theoretischen Grundlagen“ angedeutet, setzt dieses Kapitel grundlegende Kenntnisse der Wahrnehmungspsychologie voraus. Im Folgenden wird daher die Tiefenwahrnehmung in ihren monokularen und binokularen Formen näher erläutert. Ein anschauliches Beispiel hierfür ist die Wahrnehmung von Objekten entlang einer Bahnstrecke. „Schauen wir während der Fahrt aus dem Zug, rauschen nahe Objekte schnell an uns vorbei, ohne dass wir sie scharf wahrnehmen können, während Objekte am Horizont nahezu unverändert bleiben“ (Hergovich, 2022b, S. 76). Dieses Phänomen wird als Bewegungsparallaxe (*Parallax Effect*) bezeichnet und gilt als monokularer Tiefenhinweis, da bereits ein Auge genügt, um relative Tiefenunterschiede wahrzunehmen (Hergovich, 2022b, S. 76).

Im Rahmen der binokularen Tiefenwahrnehmung kommt es aufgrund der leichten Versetzung der Augen zum Auftreten von zwei verschiedenen Bildern, die in der Fachliteratur als Querdisparation bezeichnet werden. Das Gehirn gleicht diese beiden Bilder ab und unterscheidet, inwiefern sich die Positionen der Objekte voneinander versetzt sind (Schönhammer, 2013, S. 196–197). Die daraus resultierende Tiefenwahrnehmung übertrifft die eines einzelnen Auges (Hibbard et al., 2023, S. 7–8). Um die Thematik zu veranschaulichen, kann der Versuch unternommen werden, einen Faden durch ein Nadelöhr zu fädeln, zunächst mit einem geschlossenen Auge und anschließend mit beiden Augen.

### Grenzen der Tiefenwahrnehmung auf 2D-Displays

Der Versuch, Tiefe auf einem 2D-Monitor darzustellen, auf den beide Augen fokussiert sind, führt nicht zur Querdisparation, da das Bild lediglich auf einer flachen Oberfläche betrachtet wird (vgl. Lehn et al., 2022, S. 422). Die Tiefenwahrnehmung erfolgt daher ausschließlich monokular, auch wenn beide Augen genutzt werden. Um dennoch einen Eindruck von Tiefe zu erzeugen, können verschiedene Gestaltungsmittel eingesetzt werden. So lassen sich durch perspektivische Fluchtpunkte oder Größenunterschiede von Objekten räumliche Effekte hervorrufen. Auch eine gezielte Nutzung von Licht und Schatten oder Hierarchie der Objekte, wenn im Vordergrund stehende Objekte die im Hintergrund verdecken, können Tiefe im Bild schaffen. (vgl. Hergovich, 2022b, S. 77).

Darüber hinaus kann Bewegung eine Bewegungsparallaxe erzeugen und die Tiefenwahrnehmung verstärken. Auch externe Faktoren beeinflussen die Wahrnehmung: Untersuchungen zeigen, dass sowohl der Betrachtungsabstand als auch die Displaygröße eine Rolle spielen. Je näher der Betrachter am Monitor sitzt und je größer das Display ist, desto stärker wirken Fluchtpunkte und Größenunterschiede, wodurch die Tiefe intensiver wahrgenommen wird (vgl. Miyashita et al., 2022, S. 5:2–5:5). Es ist wichtig auf diese Methoden zuzugreifen, um mehr tiefe auf 2D-Monitoren zu erzeugen, denn auf Standbildern bleibt der Tiefeneindruck eher flach (Miyashita et al., 2022, S. 5:10-5:11).

## 360°-Panoramen als Darstellungsformat

Monitore und Bildformate besitzen feste Grenzen und beschränken somit den sichtbaren Bildausschnitt (Hoffmann, 2025, S. 10–11). Sie unterliegen normierten und standardisierten Größenverhältnissen, die sich historisch aus technischen Entwicklungen ergeben haben (Hoffmann, 2025, S. 19–20). Um den wahrnehmbaren Darstellungsbereich über diese Festgelegten Grenzen zu erweitern, werden Panoramen eingesetzt. Diese ermöglichen eine erweiterte bildliche Wahrnehmung sowohl auf zweidimensionalen Bildschirmen oder im VR-Bereich. (Hoffmann, 2025, S. 11–12).

Dabei kann man unter verschiedenen Arten von Panoramabildern unterscheiden, am bekanntesten sind 180°-Darstellungen. Für diese Bachelorarbeit werden jedoch **360°-Panoramen** betrachtet, da diese im Kontext von **3D-Visualisierung und Tiefenwahrnehmung** relevanter sind. (Hoffmann, 2025, S. 23–25).

Diese Darstellungsform ist besonders hilfreich, um seine Umgebung von einem Zentralen Punkt aus wahrzunehmen. In der Regel sehen wir unser Sichtfeld um uns herum, ein Panorama bildet dieses Sichtfeld auf eine Flache ebene ab und erzeugt damit eine Verzerrung (Hoffmann, 2025, S. 21–24).

### Equirectangular Projection und Geometrische Verzerrungen

Wir betrachten die Welt in einer sphärischen Geometrie unser gesamtes Sichtfeld bildet eine Kugeloberfläche 360°, jeder punkt in diesen raum lässt sich durch Richtungsvektoren beschreiben. Die reale Umgebung ist in sphärischen Koordinatenkoordinaten definiert. Sobald man diese Umgebung auf eine 2D Fläche überträgt, in Form eines Bilds oder ähnliches wird im Oberen und unterem Bereich das Bild verzerrt. Das 360°-Panorama wird an der Oberen und Unteren Bildfläche so abgewickelt, dass Breitengrade linear auf die y-Achse und Längengrade linear auf die x-Achse gemappt werden. Mathematisch erklärt:

Längengrad € [-180,180] und Breitengrad € [-90,90].

Da diese lineare Abbildung den Umfang der Kugel auf eine flache Ebene „abgewickelt“, entstehen Verzerrungen, die zum Rand hin zunehmen. Der Maßstab wächst mit dem Breitengrad gemäß (Abb. blabla)

Der Maßstab M Wächst mit der breite:

Am Äquator ist → keine Verzerrung

An den Polen geht → maximale Verzerrung

Ein Bild, das Diagramm, Kreis, Reihe enthält.

KI-generierte Inhalte können fehlerhaft sein.

* Wir machen ein vergleich
* Bei einem der 3 Varianten in diesem Vergleich, ist Verzerrung am Start
* Was ist bildliche Verzerrung?
* Lässt sie sich minimieren? Oder Verhindern
* Wenn nein was könnte man dagegen tun? Banking?



### Rolle von Depth Maps

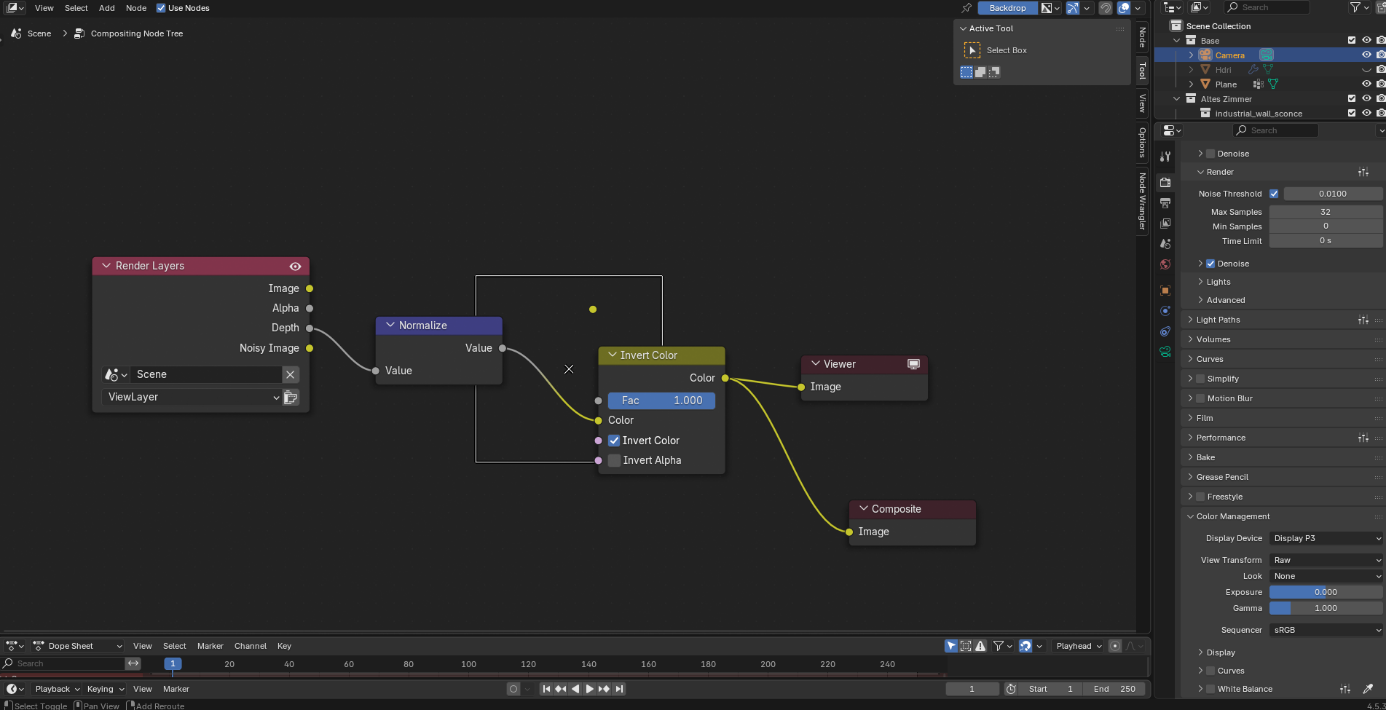
Eine Depth Map ist ein Bild, das Tiefeninformationen einer Szene speichert (Jähne, 2024, S. 271–272). Sie gehört wie eine Normal Map oder eine Bump Map zur Kategorie der Textur Maps. Gespeichert wird dabei der Abstand zwischen der Kamera und dem jeweils nächstgelegenen 3D Punkt der Szene (Jähne, 2024, S. 283–289). Die Darstellung erfolgt in der Regel über einen Farbverlauf, meist zwischen Schwarz und Weiß. Helle Bereiche stehen für nahe Objekte, während dunkle Bereiche größere Entfernungen kennzeichnen. Grundsätzlich sind auch andere Farbschemata möglich, die Monokular-Depth Graustufenabbildung gilt jedoch als Standard (Divyanth et al., 2023, S. 2–3).

Depth Maps können auf unterschiedliche Weise erzeugt werden. Zu den gängigen Methoden zählen Stereovision, Time-of-Flight, Structured Light oder monokulare Schätzungen mithilfe von KI (Jähne, 2024, S. 283–289). Im hierfür wurde Blender verwendet, wobei über die Compositing-Nodes ein sogenannter Z-Depth-Pass aus dem Render extrahiert wurde. Blender kann die Tiefeninformationen direkt aus der 3D-Szene berechnen, indem für jedes Pixel der Abstand zur Kamera entlang der Z-Achse bestimmt wird. Diese Werte werden im sogenannten Z-Buffer gespeichert und ergeben in ihrer Gesamtheit die Depth Map (Blender Foundation, 2025).

Gängige Einsatzfelder von Depth Maps sind beispielsweise die Hinderniserkennung in virtuellen Umgebungen oder die Modellierung von Szenen. Darüber hinaus werden sie im sogenannten Depth-Image-Based Rendering (DIBR) genutzt, um neue Kameraperspektiven aus vorhandenen Tiefeninformationen zu berechnen (Schmeing & Jiang, o. J., S. 279–280) (Lipski et al., 2014, S. 943–944). Auch in modernen Anwendungen, etwa im autonomen Fahren oder in der Robotik, dienen Depth Maps als Grundlage für die Erstellung topologischer Karten zur Erfassung von Hindernissen und Entfernungen (Saha et al., 2022, S. 1).

Schwierigkeiten zur Tiefeneinschätzung hat eine Depth Map mit spiegelnden Oberflächen, zudem fassen sie nur die Vorderfläche von einer Szene auf und liefern dadurch eine eingeschränkte Sicht. Außerdem sind sie häufig verrauscht und enthalten oft Schätzungsfehler, wenn man sie mithilfe einer KI erstellt (Divyanth et al., 2023, S. 6–7).

Mit all diesen Möglichkeiten kann man weiterführend viel mit Depth Maps machen, kombiniert Depth Maps aus verschiedenen Blickwinkeln können zu einer dichten 3D-Rekonstruktion verhelfen (Izadi et al., 2011, S. 1). Neuerdings kann man mit Open-Source Frameworks, wie zum Beispiel NeRF (Neural Radiance Fields), realistische 3D-Szenen mit ihnen erstellen (Ito et al., 2024, S. 5). (In dieser wird das Szenario 3 nach dem Prinzip erstellt und ist daher für die Theorie wichtig.)



## Überblick bestehender Ansätze

## Stand der Forschung im Bereich Tiefenwahrnemung

# Methodik

## Forschungsdesign: Analyse und Bewertung der Szenarien

* Wir machen ein Vergleich
* Bewertet wird sowas wie

## Definition der Szenarien

Hier erkläre ich das ich die 3 Szenarien nun beschreiben werde und wie der Aufbau umgesetzt wurde. Skizzen für jedes Szenario erstellt, um ein Verständnis darüber zu bekommen, was ich mir gedacht habe.

### Szenario 1: Baseline-3D-Rendering mit gezielten Tiefenherausforderungen

Zur Erstellung des Baseline-Szenarios wurde in Blender ein 3D-Modell entwickelt. Dabei kam das Add-on *Poly Haven Asset Browser* (*Poly Haven Asset Browser*, 2025) zum Einsatz, das eine integrierte Bibliothek hochauflösender, texturierter 3D-Modelle bereitstellt, die per Drag-and-drop in die Szene eingefügt werden können. Einige Assets wurden von der Plattform Fab heruntergeladen und in die Szene eingefügt. (*Fab*, 2025) Andere wurden selbständig modelliert und ebenfalls eingebaut. Nach der Platzierung erfolgten Optimierungen, und der Szene wurden Lichter hinzugefügt. Anschließend wurden alle Modelle zusammengeführt und als *.glb*-Datei exportiert, das Format eignet sich besonders für 3d Modelle, die man im Web-Visualisieren möchte (*glTF - Runtime 3D Asset Delivery*, 2020). Die Datei wurde anschließend auf Babylon.js eingebettet und dort dargestellt. Hierbei wurde darauf wertgelegt das die Szene den Bewertungsansprüchen entspricht.

### Szenario 2: Depth Map + 360°-Panorama (*Babylon.js*, Artefakte, Optimierungsstrategien)

* Anschließend wurde ein 360° Panorama gerendert und eine Depthmap mit Blender erstellt(wie genau das gemacht wurde kann ein eigenes Unterkapitel werden.)

### Szenario 3: Low-Poly-Rekonstruktion aus Panorama + ggf. Depth Map (Tools wie zForm, HoHoNet)

## Technische Umsetzung

## Evaluationsdesig

### Stimulusmaterial

### Auswertungsmethode

# Ergebnisse

## Technische Ergebnisse der Szenario-Umsetzungen

## Beobachtete Optimierungseffekte (z. B. weniger Artefakte, visuelle Glätte)

## Ergebnisse der Nutzerbefragung (Darstellung in Tabellen, Diagrammen)

## Zusammenfassung der Ergebnisse in Bezug auf Forschungsfrage

# Diskussion

### Interpretation der Ergebnisse im Kontext der Forschungsfrage

### Vergleich mit Literatur(?)

### Bewertung der Verfahren: Stärken, Grenzen, Anwendungspotenzial

# Fazit und Ausblick

## Beantwortung der Forschungsfrage

## Beitrag der Arbeit

## Offene Fragen und Ausblick

# Anhang