Berliner Hochschule für Technik

Immersives und interaktives AV-Streaming von Veranstaltungen

Abschlussarbeit zur Erlangung des Akademischen Grades eines

Master of Engineering (M.Eng.)

Fachbereich VIII - Maschinenbau, Veranstaltungstechnik, Verfahrenstechnik
Studiengang - Veranstaltungstechnik und -management (Master of Engineering)

Erstbetreuer: Prof. Dr. rer. nat. Alexander Lindau
Zweitbetreuer: Herr Carsten Stabenow

Vorgelegt von: Mohamed Aziz Sghaier
Matrikel-Nr.: 877036

Abgabedatum: 31. August 2022

Danksagung

Herzlichen Dank an Herrn Prof. De. rer. nat. Alexander Lindau und Herrn M.A. M. Carsten Stabenow für die hilfreiche Beratung und Begutachtung bei der Konzeption, sowie für die Dokumentation dieser Forschungsarbeit. Für die Bereitstellung der Leihmaterialien der BHT möchte ich mich bei Herrn Dipl.-Des. Arnim Bautz bedanken.

Einen besonderen Dank auch an den Theaterleiter, Herrn Stefan Meinhold, vom Bluemax Theater Berlin, sowie an den künstlerischen Leiter und selbst aktiver Blue Man, Herrn Nadim Helow, die mir ihr Vertrauen geschenkt haben und ermöglicht haben, die Blue Man Group Show aufzunehmen.

Des Weiteren möchte ich der Indie-Pop-Band THALA für ihre Teilnahme an den Aufnahmen danken, sowie meinen Kommilitonen Sebastian Finsterer, Marius Mirschberger und Mohamed Amine Mechria für ihre tatkräftige Unterstützung während der Videoaufnahmen auf dem Immergut Festival und Bluemax Theater Berlin.

Schließlich möchte ich mich bei meiner Familie bedanken, für die moralische und finanzielle Unterstützung, sowie bei den Teilnehmenden der Umfrage.

Abstract

Mit der 360°-Videoaufnahme entsteht ein Rundumbild, welches die gesamte Umgebung des Betrachters abbildet. Der Betrachter, ausgestattet mit einer VR-Brille, wird damit Teil der virtuellen Welt. Hierdurch sollen für ihn neue Sinneseindrücke geschaffen werden. Der Einsatz dieses neuen Aufzeichnungsformats in der Veranstaltungsbranche wird mit der üblichen 2D-aufnahmeformat verglichen. Zu diesem Zweck wurde eine Theatershow der Blue Man Group im Bluemax Theater Berlin parallel mit einer 2D- und einer 360°-Kamera aufgezeichnet. Anschließend wurden diese beiden Arten von Aufnahmen anhand eines Fragebogens miteinander verglichen, um herauszufinden, ob der Einsatz dieser neuen Technik empfehlenswert ist. Die Ergebnisse der Untersuchung zeigen, dass der Einsatz dieser aufkommenden Technologie im Veranstaltungsbereich für den allgemeinen Verbraucher durchaus attraktiv sein könnte.

The 360° video recording creates an all-round image that shows the viewer's entire surroundings. The viewer, equipped with VR glasses, thus becomes part of the virtual world. This is intended to create new sensory impressions for him. The use of this new recording format in the event industry is compared with the usual 2D recording format. For this purpose, a theater show of the Blue Man Group at the Bluemax Theater Berlin was recorded in parallel with a 2D and a 360° camera. Subsequently, these two types of recordings were compared with each other using a survey to find out whether the use of this emerging technology is recommendable. The results of the study show that the use of this new technology in the event sector could certainly be attractive for the general consumer.

Inhaltsverzeichnis

Danksagung II

Abstract…… III

Inhaltsverzeichnis IV

Abkürzungsverzeichnis VII

1 Einleitung 1

2 Theoretische Grundlagen 3

2.1 Streaming 3

2.1.1 Livestreaming 4

2.1.2 Livestreaming mit Interaktion 4

2.1.3 VR-Streaming Plattformen 9

3 Konventionelle 2D Videoaufnahme 12

3.1 Visuelle Wahrnehmung 12

3.2 Technische Komponenten und Grundbegriffe 13

3.2.1 Auflösung 13

3.2.2 Seitenverhältnis 14

3.2.3 Bildrate 15

3.2.4 RGB-Farbsystem 15

3.2.5 Digitalisierung des Bildsignals 16

4 Virtual Reality 19

4.1 Historische Entwicklung 19

4.2 Technische Komponenten 21

4.2.1 Stationäre VR-Brillen (HMD) 21

4.2.2 Mobile VR-Brillen 22
4.2.3 Tracking und Bewegung 22
4.2.4 Technische Eigenschaften der 360°-3D-Kameras 23
4.3 Visuelle Wahrnehmung (VR) 24
4.3.1 Präsenz und Immersion 26
4.3.2 VR-Krankheiten 28
4.4 Anwendung der VR im Veranstaltungsbereich 29
5 3D Audio 31
5.1 Auditive Wahrnehmung 31
5.2 3D Audio 32
5.2.1 Ambisonics 33
5.2.2 Tetraedermikrofone 34
5.2.3 Ambisonics erster Ordnung 35
5.3 3D Aufnahme 35
5.3.1 B-Format Konvertierung 37
5.3.2 Output-Formate 37
5.3.3 Binaural Audio über Kopfhörer 38
5.4 Plug-ins zur 3D Audioverarbeitung 40
5.4.1 IEM Plug-in Suite 40
5.4.2 Facebook Spatial Workstation 40
5.5 Anwendung der 3D Audio im Veranstaltungsbereich 41
6 Methodik 43
6.1 Forschungsfrage 43
6.1.1 Vorgehensweise und Vorüberlegungen 43
6.1.2 Inhalt des Fragebogens 46

6.2 Praktische Vorgehensweise 52

6.2.1 Durchführung einer Probe/Testaufnahme am Immergut Festival 56

6.2.2 Durführung der Echtzeit Aufnahme im Bluemax Theater 59

6.2.3 Postproduktion 65

6.2.4 Veröffentlichung 73

7 Auswertung und Interpretation der Ergebnisse 74

7.1 Allgemein 76

7.2 Ergebnisse des Vergleichs des Sinneseindrucks Aspekt 78

7.3 Ergebnisse der Bewertung des wirtschaftlichen Aspektes 83

7.4 Ergebnisse der Bewertung der Nutzerkomfort Aspekt 84

8 Fazit und Ausblick 88

Abbildungsverzeichnis 90

Literaturverzeichnis 93

Anhang……. 101

A. Einleintung zum Fragebogen 101

B. Fragebogen Seite 2 102

C. Fragebogen Seite 3 102

D. Fragebogen Seite 4 103

E. Fragebogen Seite 5 103

Abkürzungsverzeichnis

2D.....*Zweidimensional*

3D ..... Dreidimensional

3DoF ..... *Three-Degrees-of-Freedom*

5G ..... *Fünfte Generation*

6DoF ..... *Six degrees of freedom*

AAC ..... *Advanced Audio Coding*

Abb ..... *Abbildung*

ACN ..... *Ambisonic channel order*

AmbiX ..... *Ambisonics exchangeable*

App ..... *Application*

AR ..... *Augmented Reality*

AV ..... *Audiovisuel*

BHT ..... *Berliner Hochschule für Technik*

bps ..... *Bits per second*

BTA ..... *Bluetooth Adapter*

CBR.....Constant Bit Rate

CCD ..... *Charge-coupled device*

DAW..... *Digital Audio Workstations*

db.....*Decibel*

dBFs ..... *Decibels relative to full scale*

E-Learning ..... *Electronic Learning*

FOA ..... *First Order Ambisonics*

FOV ..... *Front Of View*

fps ..... *Frames Per Second*

FuMa..... *Furse-Malham*

HD ..... *High Definition*

HDTV..... *High Definition Television*

HEVC ..... *High Efficiency Video Coding*

HMD ..... *Head-Mounted Display*

HRTF ..... *Head-related transfer functions*

Hz ..... *Hertz*

Kbps ..... *Kilobits per second*

LFUS ..... *Loudness Units relative to Full Scale*

MAC ..... *Macintosh*

MB ..... *Megabyte*

Mbps ..... *Megabit per second*

MIPS ..... *Millionen von Befehlen pro Sekunde*

MP3 ..... *MPEG-1 Audio Layer III*

MP4 ..... *MPEG-4 Part 14*

MPEG ..... *Moving Pictures Experts Group*

ms ..... *Millisekunden*

NDI ..... *Network Device Interface*

PC ..... *Personal Computer*

PR ..... *Public Relation*

RGB ..... *Red, green and blue*

RTMP ..... *Real Time Messaging Protocol*

SD ..... *Secure Digital*

TCP ..... *Transmission Control Protocol*

TV ..... Television

UDP ..... *User Datagram Protocol*

UHD ..... *Ultra High Definition*

UHDTV ..... *Ultra High Definition Television*

USB ..... *Universal Serial Bus*

VBR ..... *Variable Bit Rate*

VR ..... *Virtual Reality*

VST ..... *Virtual Studio Technology*

WAV ..... *Waveform Audio File Format*

WLAN ..... *Wireless Local Area Network*

1 Einleitung

Seit Beginn der Pandemie im März 2020 bestehen umfangreiche Beschränkungen der Versammlungsfreiheit. Hierdurch kam es zu einer massiven Begrenzung von öffentlichen und nicht öffentlichen Veranstaltungen, Versammlungen, Zusammenkünften und Ansammlungen von Menschen. Die Veranstaltungsbranche wurde stark getroffen und rechnete mit Umsatzeinbußen in Milliardenhöhe sowie einem starken Abbau von Arbeitsplätzen (AlarmstufeRot, 2020).

Doch wie können Veranstaltungen trotz aller Beschränkungen weitgehend normal durchgeführt werden? Wie können Musiker*innen und Schauspieler*innen weiterhin ihrem Beruf nachgehen und auf der Bühne auftreten? Wie können mögliche Hybrid-Veranstaltungen umgesetzt werden, damit die Zuschauer sicher und gesund teilnehmen können?

In Zeiten der Digitalisierung gibt es zahlreiche Technologien und Lösungen, um Veranstaltungen direkt zu streamen bzw. im Internet zu veröffentlichen. In dieser Studie wird zwischen einer konventionellen 2D-Videoaufnahme und einer 360°-Videoaufnahme unterschieden.

Für die Umsetzung eines 360°- Immersiven Erlebnisses wurde eine 360°-Kamera verwendet (Vuze 3D-360°- 4K-VR -Kamera), diese ist mit acht Linsen ausgestattet und ermöglicht 4K stereoskopische 360°-Videoaufnahmen. Die vier internen Mikrofone der Kamera ermöglichen die Stereotonaufnahme. Zusätzlich wird für dieses Projekt ein weiteres 3D-Mikrofon (Sennheiser AMBEO VR Mic) verwendet, um eine räumliche 360°-Tonaufnahmen zu ermöglichen, die zu dem VR-Inhalt passend ist, um ein möglichst echtes immersives Erlebnis für die Zuschauer*innen zu gewährleisten. Der Inhalt ist ein kurzer Videoschnitt von einer Theatershow der Blue Man Group im Bluemax Theater Berlin (Stage Entertainment, 2022), der etwa 6 Minuten dauert. Die Show wurde parallel mit zwei Kameras (2D-Kamera und 360°-Kamera) aufgezeichnet, anschließend wurden die Videoaufnahmen im Internet veröffentlicht.

Um die Auswirkungen der Nutzung der VR-Technologie zu bewerten bzw. zu vergleichen, wurde eine Umfrage mit 15 Probanden durchgeführt, um ihre Meinung über den Vergleich der Videoformate abzubilden. Mit den erhobenen Daten wurde eine

quantitative Auswertung erarbeitet. Die dadurch entstehenden Ergebnisse wurden dann erfasst und dokumentiert, um zu ermitteln, ob der Einsatz von 360°-Videoaufnahmen im Veranstaltungsbereich gerechtfertigt ist.

Im ersten Teil der Arbeit werden die theoretischen Grundlagen erläutert. Kapitel sechs beschreibt das Vorgehen während der Durchführung des Experiments und beschreibt die praktische Umsetzung. Die Interpretation der Ergebnisse wird in Kapitel sieben behandelt. Die Arbeit schließt mit Kapitel acht und dem Fazit dieser Arbeit.

2 Theoretische Grundlagen

Das folgende Kapitel gibt einen Überblick über die theoretischen Grundlagen des Forschungsthemas. Es werden die grundlegenden Funktionen von Streaming erläutert. Danach wird erklärt, wie Menschen sogenannte konventionelle 2D-Aufnahmen wahrnehmen und wie VR- bzw. 360°-Aufnahmen die visuelle Wahrnehmung beeinflussen. Abschließend wird das Thema 3D-Audio behandelt, indem die 3D-Audio Wahrnehmung und die Bearbeitung des Audiomaterials erklärt wird.

2.1 Streaming

Streaming bietet viele verschiedene Einsatzmöglichkeiten – von der Unterhaltung bis hin zu praktischen Geschäftsanwendungen. Es kann für die Live-Übertragung verschiedene Arten von Veranstaltungen verwendet wird. Streaming bezeichnet die kontinuierliche Wiedergabe von Medieninhalten (Audio und/oder Video) im Internet in Echtzeit (Meinel und Sack, 2012). Die Wiedergabe erfolgt bereits während der Übertragung, ohne dass die*der Nutzer*in warten muss, bis der Inhalt vollständig übertragen wurde (Meinel und Sack, 2012). Gestreamte Daten werden von einem Server übermittelt und in Echtzeit von Clientanwendungen empfangen und gerendert (Furht, 2008). Diese Clientanwendungen können mit der Wiedergabe von Audio und Video beginnen, sobald genügend Daten empfangen und im Zwischenspeicher des Empfängers gespeichert wurden (Furht, 2008) (vgl. Abb.1). Dies bedeutet, dass z. B. das Abspielen von Audio-Dateien bereits während des Downloads möglich ist. (Meinel und Sack, 2014).

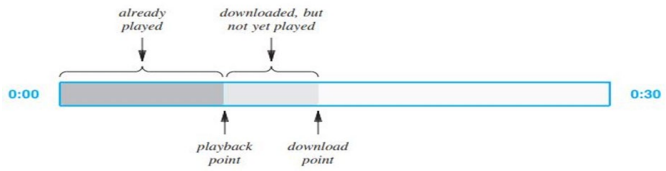


Abb. 1: Darstellung einer Anzeige, die die Zwischenspeicher für ein 30-minütiges Streaming-Video mit beschrifteten Segmenten und Punkten (Comer, 2014)

2.1.1 Livestreaming

Unter Livestreaming versteht man die Übertragung von Online-Veranstaltungen. (Heber, 2020). Das können große Konzerte, Shows und Konferenzen sein, aber auch kleinere Veranstaltungen wie E-Learning-Einheiten und Lesungen (Heber, 2020). Charakteristisch ist dabei die Live-Übertragung, d. h. die Bereitstellung des Videos in Echtzeit über das Internet (Heber, 2020). Nutzer*innen sitzen dabei einzelnen vor ihren eigenen Anzeigegeräten (PC/Laptop, mobile Geräte, VR-Brille usw.) und nicht mit anderen Personen im Publikum (Heber, 2020).

Latenz bei Livestreaming:

Die Latenzzeit im Livestream ist die Verzögerung, mit der ein Ereignis für den Zuschauer*innen angezeigt wird, nachdem es von der Kamera aufgenommen wurde. Beim Einrichten des Livestreams ist es wichtig zu berücksichtigen, wie sich die Latenz auf die Zuschauer auswirkt (YouTube-Hilfe, 2022). Eine niedrige Latenz sorgt dafür, dass man schnell mit den Zuschauern*innen in Chat und Kommentaren interagieren kann (YouTube-Hilfe, 2022).

2.1.2 Livestreaming mit Interaktion

Interaktion zwischen Teilnehmern*innen ist in konventionellen, live gestreamten Veranstaltungen relativ einfach einzusetzen. Dies erfolgt über die integrierte Chatfunktion, die das Senden und Empfangen von Nachrichten in Echtzeit ermöglicht, oder im Falle von Videokonferenzen, können sich die Teilnehmer*innen einfach sehen und in einem Live-Meeting miteinander visuell und verbal kommunizieren. Live-Streaming-Plattformen wie, Facebook (Meta, 2022) und YouTube bieten mittlerweile diese Funktion kostenlos an (Youtube-Hilfe, 2022). Für die Interaktion mehrerer Nutzer sind die Verzögerungen entscheidend. Bei Sprache werden Verzögerungen bis zu 150 ms (Millisekunden) nicht wahrgenommen, 150 ms bis 400 ms werden als akzeptabel empfunden, während Verzögerungen über 400 ms als unangenehme Störung wahrgenommen werden (Meinel und Sack 2012).

Encoder

Voraussetzung für das Streaming ist die Konvertierung (Encodierung) des Videosignals der angeschlossenen Kamera in ein Web-kompatibles Format (Heber 2020). Hierfür werden sogenannte Encoder eingesetzt (Heber 2020). Der Encoder wandelt die rohen Audio- und Videosignale in ein Format um, sodass sie über geeignete Plattformen gestreamt werden können (Mack und Rayburn, 2006). Dies ist ein erforderlicher Prozess, da die Datenraten von Roh-AV Signale ansonsten viel zu hoch für die Übertragung über das Netzwerk sind (Mack und Rayburn, 2006). Die Enkodierung erfolgt in der Regel entweder mit einem Software- oder Hardware-Encoder (Mack und Rayburn, 2006).

Hardware und Sofware Encoder

Einige Software-Encoder sind kostenlos erhältlich, wie z. B. OBS, FFSplit oder Streamlabs. Die Softwarelösungen bieten in der Regel mehr Flexibilität, da der Bildmischer enthalten ist. Daher sind sie kostengünstiger als Hardwarelösungen und bieten mehr und bieten mehr Anpassungsmöglichkeiten (Heber, 2020). Auf der anderen Seite sind Softwarelösungen von der Rechenleistung des verfügbaren Computers abhängig und haben eine höhere Latenz (Heber, 2020). Vorhandene Rechner können mit Software und Capture-Karten ausgestattet und zu Encodern umgewandelt werden (Heber, 2020). Diese Lösung ist mit den Anforderungen eines billigen Hardware-Encoders vergleichbar (Heber, 2020).

Hardware-Encoder müssen die Streaming-Anforderungen erfüllen und müssen separat angeschafft werden (Kapser, 2021). Sie bieten mehr Zuverlässigkeit, da sie unabhängig von der Rechenleistung des Computers sind, eine geringere Latenz und eine höhere Encodierungsgeschwindigkeit, sind aber wesentlich teurer und bieten weniger Anpassungsmöglichkeiten (Kapser, 2021). Die Preise für Hardware-Encoder reichen von einigen Hundert bis zu mehreren Tausend Euro (Heber, 2020). Welche Encoder Lösung Richtige ist, hängt vom jeweiligen Projekt ab (Heber, 2020).

Im Folgenden wird die OBS Studio als Softwarelösung vorgestellt.

OBS Studio

Open Broadcaster Software (OBS) wurde 2012 eingeführt. Es handelt sich um eine kostenlose und Open-Source-Software für die Aufnahme und das Streaming von Live-Video. OBS ist in C und C++ programmiert (Richards, 2019). Die Software bietet Nutzern*innen eine effiziente Alternative zu kostenpflichtiger Software (Richards, 2019). OBS Studio ermöglicht Echtzeitaufnahmen, Bildmischung, Enkodierung und Broadcasting über das Real Time Messaging Protocol (RTMP) (Wikipedia, 2022). Es kann Videos an jedes RTMP-unterstützende Ziel streamen, einschließlich YouTube, Twitch, Instagram und Facebook (Wikipedia, 2022). Das Programm unterstützt Surround-Sound und Ambisonic erster Ordnung für Facebook Live 360°, höhere Ambisonic Ordnungen werden nicht unterstützt (OBS, 2022).

Allgemeine Streaming-Audio- und Video-Codecs

Für Online-Audio und -Video werden in der Regel mehrere Codec-Typen verwendet. Einige gängige Audio- und Videocodecs sind hier in diesem Abschnitt vorgestellt:

H.264: Ist heute der am häufigsten verwendete Videocodec. Er deckt ein breites Spektrum von Einsatzmöglichkeiten ab, von hoher Bitrate bis zu sehr niedriger Bitrate. Ist zum Standard für verschiedene Anwendungszwecke, z. B. für Blu-ray-Discs, HDTV-Übertragungen, Videostreaming im Internet und Anwendungen auf mobilen und tragbaren Geräten eingesetzt (Li et al., 2014).

AAC ist derzeit der am häufigsten verwendete Audiocodec. Der MPEG-2 Advanced Audio Coding (AAC)-Standard, wurde als weiteres für Komprimierung- und Kodierungsverfahren digitale Audiodaten als Nachfolger von MP3 entwickelt und bietet bei gleicher Bitrate eine bessere Klangqualität als MP3 (Li et al., 2014). AAC ist derzeit das Standard-Audioformat für YouTube, iPhone und andere Apple-Produkte sowie für iTunes, Nintendo und PlayStation. Es wird auch von Android-Mobilgeräten unterstützt (Li et al., 2014).

MP3 ist ein beliebter Standard für die Audiokomprimierung. Die „3“ steht für den Layer 3, und „MP“ steht für den MPEG-Standard. Er wurde 1992 veröffentlicht (Li et al., 2014).

Nächste Generation von Streaming-Codecs

HEVC, oder H.265 ist der Nachfolger von H.264. Wurde im Januar 2013 veröffentlicht. Der Videocodec bietet eine höhere Videoauflösung von bis zu 8k bzw. 4k in UHDTV und eine 50% niedrigere komprimierte Videogröße (bei gleicher visueller Qualität) als H.264 (Li et al., 2014).

Bandbreite

Ist das Maß für der Übertragungsgeschwindigkeit über digitale Verbindungen oder Netzwerke, oft in Kilobits pro Sekunde (Kbps) oder Megabits pro Sekunde (Mbps) (Li et al., 2014). Je größer die Bandbreite ist, desto mehr Informationen können in einer Sekunde übertragen werden (Meinel und Sack, 2012). Die Bandbreite wird im Allgemeinen in Bits pro Sekunde (bps) gemessen. Die Videoqualität muss für das Seherlebnis gewährleistet sein, indem mehrere Bitraten auf einmal gestreamt werden, dafür wird ein Transcoder verwendet, der den Stream in verschiedenen Auflösungen zur Verfügung stellt (Heber, 2020). Dadurch werden die Inhalte einem größeren Benutzerkreis zugänglich gemacht, da der Stream auch bei einer eher langsamen Internetverbindung zugänglich bleiben soll (Heber, 2020).

Videoauflösung	Empfohlene kontinuierliche Geschwindigkeit
4K	20 Mbit/s
HD 1080p	5 Mbit/s
HD 720p	2,5 Mbit/s
SD 480p	1,1 Mbit/s
SD 360p	0,7 Mbit/s

Abb. 2: Empfohlene kontinuierliche Geschwindigkeit (YouTube-Hilfe, 2022)

Bitraten-Codierung

Man unterscheidet zwischen „Variable Bit Rate“ (VBR) und „Constant Bit Rate“ (CBR): Bei VBR entscheidet der Encoder, wie viele MB (Megabytes) für die Komprimierung eines Bildes benötigt werden (Heber, 2020). Wenn sich das Videobild nur leicht verändert, kann die Bitrate ohne Qualitätseinbußen verringert werden (Heber, 2020).

Bei CBR wird immer die gleiche Bitrate für die Kodierung verwendet (Heber, 2020). Plötzliche Spitzenwerte werden so vermieden, aber es wird manchmal unnötiger Datenverkehr erzeugt. Die Empfehlungen hängen von der jeweiligen Plattform ab (Heber, 2020). Wenn keine Empfehlung ausgesprochen wird, empfiehlt sich ein Test mit geeigneten Inhalten (Heber, 2020).

Internet-Transport-Protokolle

TCP – Transmission Control Protocol

Transmission Control Protocol Kurz (TCP) ist ein Datentransportprotokoll, der eine Verbindung zum gewünschten Kommunikation-Partner aufbaut und nach erfolgreicher Datenübertragung wieder abbaut (Meinel und Sack, 2012). TCP bietet einen zuverlässigen Transportdienst zum Versenden von Nachrichten über das Internet (Meinel und Sack, 2012).

UDP – User Datagram Protocol

User Datagram Protocol kurz UDP ist ein einfacheren verbindungslosen Datenübertragungsprotokoll, die es ermöglicht jederzeit Nachrichten an jedes beliebige Ziel zu versenden, ohne vorher erst für den Aufbau einer Verbindung sorgen zu müssen (Meinel und Sack, 2012). Im Gegensatz zu TCP ein verbindungsloser, ist das Protokoll einen unzuverlässigen Transportdienst (Meinel und Sack, 2012).

Meistverwendeten Streaming-Protokolle

Protokolle bestimmen, wie Streaming-Komponenten miteinander interagieren, eine Reihe von standardisierten Protokollen werden von Streaming-Servern verwendet,

unter denen das am gewöhnlichsten verwendet Streaming-Protokoll für die Übertragung von Videosignalen, das Real-Time Messaging Protocol - abgekürzt RTMP (Heber 2020).

RTMP

RTMP ermöglicht dem Nutzer, einen Stream vom Server anzufordern. Der Adobe Flash Media, obwohl er abgeschafft ist, wird noch für das RTMP verwendet, und kann Server in verschiedene Qualitätsstufen bereitstellen, je nachdem, ob die Netzwerkbandbreite ausreichend ist oder nicht (Castr, 2022). Wenn die Netzwerkgeschwindigkeit für eine bestimmte Videoeinstellung nicht ausreichend ist, schaltet das Echtzeit-Nachrichtenprotokoll automatisch auf eine niedrigere Videoauflösung und Bildrate um (Castr, 2022). Das RTMP-Live-Streaming-Protokoll kann sich auch an unterschiedliche Netzwerkbedingungen anpassen (Castr, 2022). Wenn die Bandbreite sinkt, wird die Qualität des Streams automatisch reduziert, bis die Zuschauer einen ausreichenden Puffer erhalten (Castr, 2022). Dank dieser Flexibilität können Videostreams in Echtzeit für verschiedene Geräte und Internetverbindungen anpassen (Castr, 2022).

Hier sind weitere wichtige Streaming Protokolle präsentiert:

NDI - Network Device Interface

NDI steht für Network Device Interface (NDI) und ist eine von NewTek entwickelte Video-over-IP-Technologie, es ist ein Open-Source-Protokoll (Newtek, 2022). Installiert in einem Ethernet-Netzwerk, sind NDI-Geräte sofort als Video- und Audio-Quelle oder als Ziel verfügbar (Newtek, 2022). NDI-Videos ermöglichen eine hohe Auflösung von bis zu 4K UHD und mehr sowie eine niedrige Bandbreite mit einer minimalen Latenz, sodass sie sich gut für den Einsatz in Live-Produktionen eignen (Newtek, 2022).

2.1.3 VR-Streaming Plattformen

Man unterscheidet zwischen WebVR-Plattformen und VR-spezifische Plattformen (Feyder und Rath-Wiggins, 2018).

Im Folgenden wird ein Überblick über die am häufigsten genutzten VR-Streaming-Plattformen dargestellt.

WebVR-Plattformen

VR-Anwendungen können über Browser genutzt werden. D. h. VR-Erlebnisse funktionieren - sowohl stationär als auch mobil - direkt über Browser wie Chrome, Safari oder Firefox (Feyder und Rath-Wiggins, 2018). WebVR-Anwendungen können somit über verschiedene Endgeräte laufen und haben ein hohes Reichweitenpotenzial (Feyder und Rath-Wiggins, 2018).

YouTube-VR

YouTube-VR ist die Erweiterung des YouTube-Streamingdienstes, mit dem die Nutzer*innen VR-Inhalte auf verschiedenen VR-Geräten aufrufen können (YouTube VR, 2022). Nutzer*innen können über die Plattform auch ganz einfach 360°- und 180°-VR-Videos hochladen und veröffentlichen. Diese WebVR-Plattform wurde für die Veröffentlichung des 360°-Videos des Masterprojekts gewählt, weil sie benutzerfreundlich, unkompliziert und über fast alle Geräte zugänglich ist.

Facebook Live 360°

Das Sozialnetzwerk bietet die Möglichkeit, 360°-Videoinhalte zu veröffentlichen (Feyder und Rath-Wiggins, 2018). Diese werden dann in der Timeline angezeigt und können ebenfalls über die mobile Facebook App angesehen werden (Feyder und Rath-Wiggins, 2018).

VR-spezifische Plattformen

Mixed Reality, SteamVR und Meta Quest Store werden hier als drei der derzeit prominentesten Plattformen aufgelistet: Die drei Plattformen bieten eine breite Palette von VR-Erlebnissen (Spiele, Videos, usw.) für VR-Brillen. Da sie nur für stationäre Geräte gedacht sind, können die Nutzer auf diesen Plattformen hochwertige VR-Erlebnisse genießen (Feyder und Rath-Wiggins, 2018).

Windows Mixed Reality (WMR)

ist eine von Microsoft entwickelte Plattform für virtuelle (VR) und Augmented-Reality-Anwendungen (AR), die im Jahr 2015 unter dem Namen Windows Holographic veröffentlicht wurde (Wikipedia, 2021).

SteamVR

SteamVR ist eine von Valve entwickelte Virtual-Reality-Hardware- und Software-Plattform, die sich auf die Implementierung von „raumfüllenden" (Room-Scale) VR-Anwendungen konzentriert, die externe Basisstationen zur Positionierung verwenden (Wikipedia, 2022). SteamVR wurde erstmals 2014 für das Oculus Rift-Headset eingeführt und später erweitert, um andere Virtual-Reality-Headsets wie das HTC Vive und Valve Index zu unterstützen. SteamVR war ursprünglich für Windows, macOS und Linux verfügbar, aber die Unterstützung für macOS wurde im Mai 2020 eingestellt (Wikipedia, 2022).

3 Konventionelle 2D Videoaufnahme

Die Begriffe „Video“ (lateinisch: ich sehe) und „Film“ beziehen sich auf illusionäre Medien, die dazu verwendet werden können, bei den Menschen das Gefühl zu erwecken, dass ihnen eine optische Realität vorgeführt wird. (Schmidt, 2012). Ende des 19. Jahrhunderts, am 18. Dezember 1895, präsentierten die Brüder Lumière die erste Filmprojektion, die heute als die Geburtsstunde des Mediums Film gilt (Schmidt, 2012). Diese Filmprojektion wurde durch den von ihnen erfundenen Cinematograph ermöglicht, der gleichzeitig als Kamera und Projektor einsetzbar war (Schmidt, 2012). Einige Jahrzehnte später, im Laufe der 30er Jahren, löste der Tonfilm die Stummfilme ab und markierte das Ende einer Film-Ära (Schmidt, 2012). Die Kolorierung von Schwarz-Weiß-Bildern mit dem Technicolor-Verfahren wurde erst in den 40er Jahren eingeführt und revolutionierte, insbesondere im Fernsehbereich (Schmidt, 2012). Die Entwicklung der Videotechnik wurde dann vorangetrieben bis zur Einführung der Digitaltechnik im Bereich Ton und Bild Anfang der 90er Jahre, ermöglicht durch den Einsatz von leistungsfähigeren Computern, die längere Filmsequenzen in höherer Auflösung verarbeiten konnten (Schmidt, 2012). Das Aufkommen hochauflösender Digitalkameras zu Beginn des 21. Jahrhunderts war schließlich der letzte Meilenstein der heutigen digitalen Medientechnologie, die dank der Begeisterung des Publikums und des Strebens nach Innovation, die 2D-Videoaufnahme in der modernen Kultur fest etabliert hat (Schmidt, 2012).

3.1 Visuelle Wahrnehmung

Das menschliche Sehen beruht auf der Umwandlung von Lichtenergie in Nervenreize, die dann vom Gehirn verarbeitet wird (Schmidt, 2012). Die Abbildung. 3 zeigt einen Querschnitt des menschlichen Auges (Schmidt, 2012). Das Licht dringt durch die Hornhaut und die Bindehaut und fällt auf die Pupille, die wie eine Blende wirkt und vom Lichteinfall abhängig ist (Schmidt, 2012). Dies ermöglicht die Anpassung (Adaptation) an unterschiedliche Helligkeiten (Schmidt, 2012). Das Licht gelangt dann durch die Linse, deren Form mithilfe von Muskeln veränderbar ist, sodass unterschiedlich weit entfernte Objekte scharf auf der Netzhaut abgebildet werden können (Akkommodation) (Schmidt, 2012). Die Abbildungseigenschaft des Auges ergibt sich aus der Linse und der Krümmung

der Hornhaut (Schmidt, 2012). Die Netzhaut besteht aus etwa 120 Millionen Stäbchen und nur ca. 7 Millionen Zapfen (siehe Abb.4). Nur die Zapfen sind farbempfindlich, und die große Mehrheit von ihnen befindet sich in der Fovea centralis, dem Sehzentrum des Auges (Bühler et al., 2018). Es gibt drei verschiedene Arten von Zapfen, nämlich Rot-, Grün- und Blaurezeptoren, die bevorzugt rotes, grünes bzw. blaues Licht absorbieren (Bühler et al., 2018). Wir sehen also nur drei Farben: rot, grün und blau (Bühler et al., 2018). Alle anderen Farben sind das Ergebnis der Signalverarbeitung und -auswertung im Sehzentrum des Gehirns (Bühler et al., 2018). Die Stäbchen sind sehr empfindlich und registrieren schwache Helligkeitswerte (Schmidt, 2012). Der Mensch nimmt also seine Umgebung wahr, indem er Objekte permanent mit der Fovea centralis abtastet (Schmidt, 2012). Während des Abtastens passt sich das Auge ständig an die neue Blickrichtung an (Schmidt, 2012). Der visuelle Gesamteindruck ergibt sich dann aus dieser Abtastung und der Gedächtnisleistung des Gehirns (Schmidt, 2012).

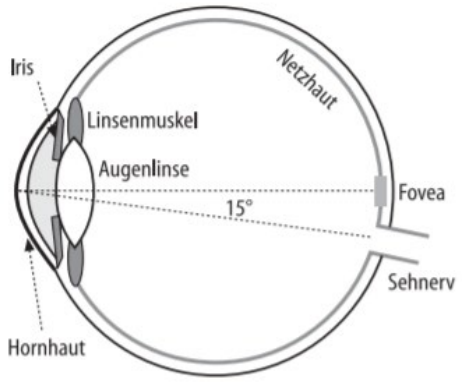


Abb. 3: Querschnitt des menschlichen Auges (Schmidt, 2012)

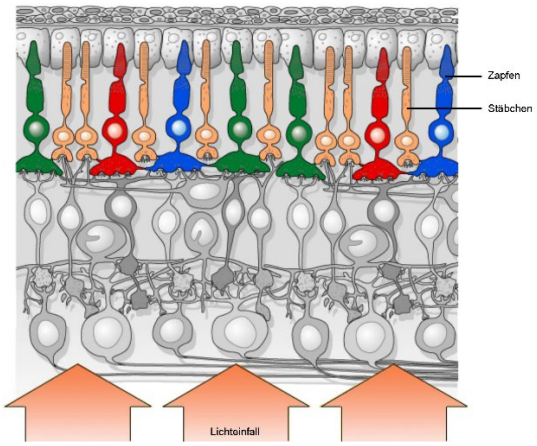


Abb. 4: Bau der Netzhaut, Zapfen und Stäbchen (Klett, 2014)

3.2 Technische Komponenten und Grundbegriffe

3.2.1 Auflösung

Die Anzahl der Pixel in der Breite und Höhe des Bildes wird als Auflösung bezeichnet (Bühler et al., 2018). Dies bedeutet, dass je höher die Pixelanzahl ist, desto detaillierter und schärfer erscheint das Bild (Bühler et al., 2018). In der folgenden Tabelle (siehe Abb. 5) ist eine Übersicht über die aktuellen Videostandards zu sehen (Bühler et al.,

2018). In dieser Arbeit wird die 4K-Auflösung für die Videoaufnahmen verwendet. 3.840 x 2.160 px, K= 1000 (abgerundet) der Anzahl der Pixel in der Bildbreite (Bühler et al., 2018).

Name	HD	Full HD (2K)	UHD-1 (4K)	UHD-2 (8K)
Auflösung (in Pixel)	1.280 · 720 px	1.920 · 1.080 px	3.840 · 2.160 px	7.680 · 4.320 px
Seitenverhältnis (Aspect Ratio)	16:9	16:9	16:9	16:9
Bildrate (in Bilder/s) i: interlaced (Halbbilder) p: progressive (Vollbilder)	25p 50i 50p 60p 100p 120p	25p 50i 50p 60p 100p 120p	50p 60p 100p 120p	50p 60p 100p 120p
Farb(unter)abtastung	4:2:2 4:2:0	4:2:2 4:2:0	4:4:4 4:2:2	4:4:4 4:2:2

Abb. 5: Aktuelle Videostandards (Bühler et al., 2018)

3.2.2 Seitenverhältnis

Das Seitenverhältnis beschreibt das Verhältnis zwischen der Breite und der Höhe eines Bildes. Man unterscheidet diesbezüglich folgende Formate:

- 4:3 TV-Format, das Bild hat also ein Seitenverhältnis von 1,33:1 (Breite: Höhe) (Bühler et al., 2018).
- 16:9 HDTV und UHDTV-Format, entspricht einem Seitenverhältnis von 1,78:1. näher an den physikalischen Eigenschaften des Auges, da das horizontale Sichtfeld des menschlichen Sehens einen Winkel von fast 180° aufweist, während es in der Vertikalen nur 30° beträgt (Bühler et al., 2018).
- 21:9 Kinoformat entspricht einem Seitenverhältnis bis zu 2,35:1 (Bühler et al., 2018).

Abb. 6: Beispiel Seitenverhältnis (eigene Darstellung)

3.2.3 Bildrate

Die Bildrate beschreibt die Anzahl der Bilder pro Sekunde, dies wird auch Bildwiederholungsfrequenz benannt (Bühler et al., 2018). Das in den Ursprüngen des Fernsehens verwendete Interlace-Verfahren, bei dem mit Halbbildern gearbeitet wurde, gibt es nur noch bei der HD bzw. Full HD (50i) (Bühler et al., 2018). Die Abkürzung lautet 50i (i für interlaced). Computerbildschirme arbeiten im Gegensatz zum Fernsehen nur mit Vollbildern - man spricht dann von progressiven Bildern (Abkürzung=P) (Bühler et al., 2018). Die Angabe 25p bedeutet also, dass 25 Vollbilder pro Sekunde gezeigt werden (Bühler et al., 2018)

3.2.4 RGB-Farbsystem

Digitalkameras sind so konzipiert, dass sie dem menschlichen Auge bestmöglich ähneln (Bühler et al., 2018). Sie bestehen aus einer Optik (vergleichbar mit einer Linse) und einem CCD-Element (vergleichbar mit Zapfen), die das Licht in seine Grundfarben Rot, Grün und Blau aufspalten, digitalisieren und speichern (Bühler et al., 2018). Jeder Pixel eines Bildschirms wird von drei Lichtquellen - Rot, Grün und Blau – angesteuert (Bühler et al., 2018). Die Farbe ergibt sich dann durch die Kombination der drei Farbkomponenten, und dies wird als RGB-Farbsystem (siehe Abb.7) bezeichnet (Bühler et al., 2018).

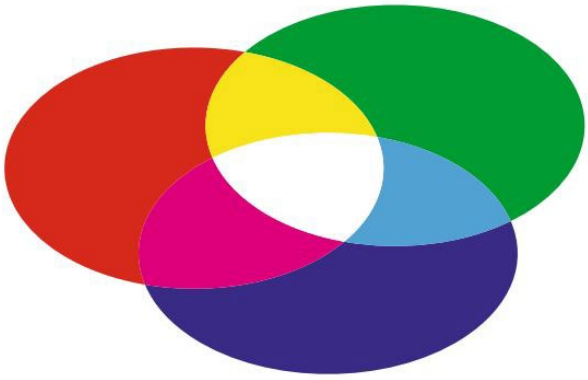


Abb. 7: Additive Farbmischung Drei Lichtpunkte der Primärfarben RGB (Böhringer et al., 2014)

Um die riesige Datenmenge der Bewegtbilder zu reduzieren, ohne die Bildqualität zu verschlechtern, wurde das RGB-System in das YCrCb-System umgewandelt, dieses System bezieht sich auf die Helligkeit und den Farbanteil, denn das Auge kann Helligkeit viel besser unterscheiden als Farbe (Bühler et al., 2018).

- Y: Helligkeit
- Cb: blau-gelber Farbanteil
- Cr: rot-grüner Farbanteil

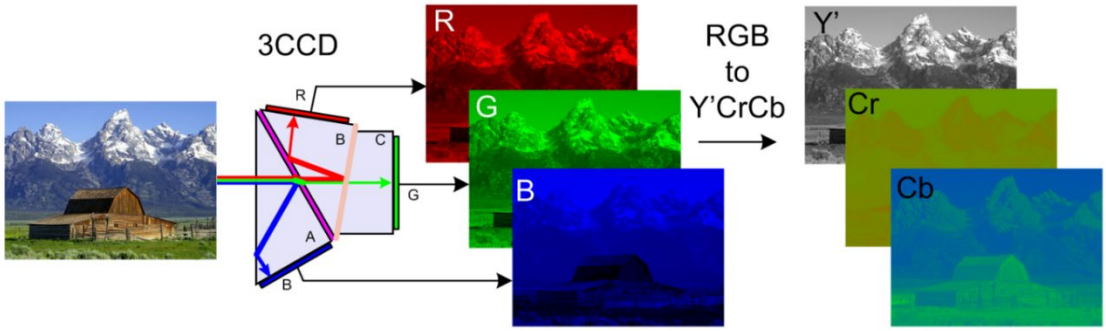


Abb. 8: Beispiel der Umwandung von RGB zuYCrCb (Wikipedia, 2012)

3.2.5 Digitalisierung des Bildsignals

Im ersten Schritt der Digitalisierung von Bildern wird das Bildsignal abgetastet (Bühler et al., 2018). Im zweiten Schritt werden die Messwerte in Binärdaten umgewandelt (Bühler et al., 2018). Bei diesem Prozess entstehen große Datenmengen und Raten, die eine Übertragung des digitalen Signals in Echtzeit unmöglich machen (Bühler et al., 2018). Um die Datenrate zu reduzieren, werden weitere Maßnahmen getroffen. Insbesondere erfolgt dies durch die Farbverarbeitung (Bühler et al., 2018).

Farbunterabtastung

Bei der Farbunterabtastungsmethode ist der Helligkeitsanteil wichtiger als die beiden Farbanteile (Bühler et al., 2018). In der oben beschriebenen Umwandlungsmethode von RGB in YCbCr wird dann die Helligkeit Y von den beiden Farbanteilen Cb und Cr getrennt (Bühler et al., 2018). Während der Digitalisierungsverfahren werden, dann die beiden

Farbanteile in einem Farbwert zusammengefasst, um die Datenmenge zu verringern (Bühler et al., 2018).

Dabei lassen sich drei Möglichkeiten unterscheiden:

- 4 : 4 : 4-Digitalsignal (ohne Datenkompression). Dadurch wird die optimale Bildqualität mit einer maximalen Farbtiefe von 24 Bit/Pixel erzeugt (Bühler et al., 2018).
- 4 : 2 : 2-Digitalsignal (mit Datenkompression). Die Datenmenge reduziert sich hierdurch auf 8 + 4 + 4 = 16 Bit/Pixel, verringert sich also um 33% (Bühler et al., 2018).
- 4 : 1 : 1-Digitalsignal (mit noch stärkere Datenkompression). Die Datenmenge reduziert sich hierdurch auf 8 + 2 + 2 = 12 Bit/Pixel, verringert sich also um 50% (Bühler et al., 2018).

Video-Codierung

Um ein Video von einer Quelle (Kamera) zu seinem Ziel (Bildschirm) zu übertragen, sind eine Reihe von Prozessen erforderlich (Bühler et al., 2018). Die wichtigsten Schritte bei diesem Prozess sind die Kompression (Codierung) und die Dekompression (Decodierung), bei denen die Rohvideoquelle auf eine optimale Größe, für die Übertragung oder Speicherung reduziert und anschließend für die Wiedergabe rekonstruiert wird (Bühler et al., 2018). Der Komprimierungsprozess erfordert die Anwendung eines Algorithmus auf das Originalvideo, um eine komprimierte Datei zu erstellen, die zur Übertragung oder Speicherung geeignet ist (Bühler et al., 2018). Ein umgekehrter Algorithmus wird auf das komprimierte Video angewendet, um ein Video zu erzeugen, das nahezu die gleiche visuelle Qualität wie das Originalvideo aufweist (Bühler et al., 2018).

MPEG-4/ H.264

Das Encoder MPEG-4, das von der MPEG (Moving Picture Experts Group) entwickelt wurde, und der Videocodec H.264 sind derzeit die wichtigsten Verfahren zur Videokompression (Bühler et al., 2018). MPEG-4 erhielt Anfang 2000 den Status einer offiziellen internationalen Norm (Li et al., 2014). Damit wird ein breites Spektrum von

Anwendungen abgedeckt, das von Überwachungskameras mit geringer Qualität und Auflösung bis zu hochauflösende Fernsehübertragungen reicht (Fischer, 2020). Der H.264 ermöglicht hohe Kodiereffizienz und Kodierflexibilität im Vergleich zu früheren Videokodierstandards (Fischer, 2020).

4 Virtual Reality

Der Begriff „Virtual Reality“ abgekürzt (VR) wurde ursprünglich von dem Leiter der Firma „Virtual Programming Language Research“ Jaron Lanier verwendet, und bezieht sich in der Regel auf eine computergesteuerte Simulation, die dreidimensional, multisensorisch und interaktiv ist, sodass der Benutzer das Gefühl hat, in einer externen Umgebung zu interagieren (Weiss et al., 2006). Die Illusion der Mitwirkung in einer virtuellen Welt basiert auf dreidimensionalen, stereoskopischen, kopfgesteuerten Displays, Körper-Tracking und binauralem Sound (Weiss et al., 2006). Ein VR-Erlebnis kann durch die Einbeziehung von, Haptik und Gerüchen erweitert und erheblich bereichert werden (Weiss et al., 2006). Eines der Hauptziele der virtuellen Realität ist es, das Aussehen der realen Welt so zu reproduzieren, dass es von der Realität nicht mehr zu unterscheiden ist (Magnor und Sorkine-Hornung, 2020).

4.1 Historische Entwicklung

Das Konzept der virtuellen Realität und die Idee, eine engere Verbindung zwischen Menschen und Computer zu schaffen, stammen aus den 60er Jahren in den Vereinigten Staaten (Warnecke et al.,1993). Ivan E. Sutherland stellte im Jahr 1965 seine Ideen der „Ultimate Display" und des Auftretens neuer Eingabe-/Ausgabegeräte in einer simulierten Computerwelt vor, aber es sollte noch einige Jahre dauern, bis Computer für solche Simulationen ausreichend entwickelt waren (Warnecke et al.,1993). Viele teure Entwicklungen von Computersimulationen mit fortschrittlichen Geräten wurden zunächst im militärischen Bereich durchgeführt, um z. B. realistische Flugsimulation zu entwickeln (Warnecke et al.,1993).

Im Jahr 1968 entwickelte Sutherland ein Head-Mounted Display System, das aus einem Datenhelm und einem mechanischen und ultraschallbasierten Tracking-System (vgl. Abb.9) (Doerner et al., 2022). Damit konnte der*die Betrachter*in eine simulierte, 3D-Umgebung in der richtigen Perspektive sehen (Doerner et al., 2022). In den 80er Jahren wurde die Firma VPL Research von Jaron Lanier und Thomas Zimmermann gegründet (Doerner et al., 2022). Das kalifornische Unternehmen war das erste Unternehmen, das Virtual-Reality-Produkte auf den Markt brachte, wie z. B. den „EyePhone"-Datenhelm, eine Weiterentwicklung des „Head Mounted Display“ von Sutherland aus den 1960er

Jahren, der eine Auflösung von 720×480 Pixeln bot (Doerner et al., 2022). Seit den 1990er Jahren wurde in diesem Bereich viel erforscht, und viele Unternehmen, insbesondere in den USA bzw. Silicon Valley, arbeiteten an der Technologie der virtuellen Realität (Doerner et al., 2022). Leider waren die PCs damals noch nicht so leistungsfähig wie heute (Doerner et al., 2022).

Während Anfang der 90er Jahre ein PC, 12 MIPS (Millionen von Befehlen pro Sekunde) ausführen konnte, ist heute jedes Smartphone in der Lage, 20.000 MIPS auszuführen (Knoll, 2017). Dank dieser Rechenleistung und den Fortschritten in der Sensortechnik konnten die Probleme der frühen Virtual-Reality-Systeme weitgehend gelöst werden (Knoll, 2017).

Abb. 9: Ivan Sutherlands Head-Mounted Display (Jewst, 2021)

Die neue Ära der virtuellen Realität begann im Jahr 2016, als die neue Generation von kopfgetragenen VR-Geräten (HMDs) mit der Oculus Rift und der HTC Vive auf den Markt gebracht wurden (Hillmann, 2019). Die Oculus VR-Brille war die erste konsumentenfreundliche Generation dieser Gerätekategorie, da sie das beste Preis-Leistungs-Verhältnis gebracht hat (Hillmann, 2019).

Mit der Pandemie bedingte Beschränkungen im Jahr 2020 erlebt diese Technologie einen weiteren enormen Sprung. Laut Grand View Research wird der globale VR-Markt bis 2027 auf 62,1 Milliarden Dollar anwachsen (Grand View Research, 2020). Darüber hinaus wird das Aufkommen der Netzwerktechnologie der fünften Generation (5G) das VR-Erlebnis aufgrund der geringen Latenz, der verbesserten Zuverlässigkeit und der hohen Dichte, die die 5G-Technologie bietet, weiter verbessern (Grand View Research, 2020). Alle Indikatoren sprechen dafür, dass die virtuelle Realität eine vielversprechende Zukunft vor sich haben wird (Grand View Research, 2020).

Am 28. Oktober 2021 wurde das Meta Plattform in Facebook vorgestellt (Meta, 2021). Laut der PR-Kampagne von Facebook handelt es sich um eine digitale Erweiterung der physischen Welt durch soziale Medien, Virtual Reality und Augmented Reality (Meta, 2021). Gleichzeitig stellte Facebook-Gründer Mark Zuckerberg eine neue VR-Brille vor, die Oculus Quest 2, ein drahtloses HMD mit 1.832 x 1.920 Pixeln pro Auge und 6-GB-Prozessor, erhältlich für etwa 350€ (Meta, 2021).

Mit all diesen technologischen Fortschritten bewegen wir uns auf eine mögliche Demokratisierung der Nutzung immersiver Technologien zu, die sicherlich viele neue Möglichkeiten und Ansätze in der heutigen Welt eröffnen wird.

4.2 Technische Komponenten

4.2.1 Stationäre VR-Brillen (HMD)

Die Stationäre VR-Brillen engl. Head-Mounted-Displays, kurz HMD oder Virtual-Reality-Brillen (vgl. Abb.10) gelten allgemein als persönliche Displays, die auf dem Kopf direkt vor den Augen der*des Nutzers*in getragen werden können (Dörner et al., 2019). Je nach Design und Gewicht werden sie wie eine Brille oder eher wie ein Fahrradhelm getragen (Dörner et al., 2019). HMDs sind häufig mit einem Trackingsystem ausgestattet, um die Blickrichtung und -position der virtuellen Kamera kontinuierlich an die aktuelle Position und Ausrichtung des Kopfes anzupassen (Dörner et al., 2019). HMDs verwenden in der Regel eine binokulare Optik, sodass die*der Nutzer*in den Inhalt stereoskopisch wahrnehmen kann (Dörner et al., 2019).

Abb. 10: Stationäre HP Reverb G2 VR-Brille (HP-Webseite)

Abb. 11: Mobile Google Cardboard VR-Brille (Feyder und Rath-Wiggins, 2018)

4.2.2 Mobile VR-Brillen

Zu diesem Zweck werden VR-Brillen wie das Google Cardboard, die Google Daydream oder die Samsung Gear VR angeboten (Feyder und Rath-Wiggins, 2018). VR-Erlebnisse können dann über Handy-Browser, Apps oder andere Plattformen wie YouTube-VR oder Facebook angeschaut werden (Feyder und Rath-Wiggins, 2018). Als Beispiel: Ein Cardboard (siehe Abb. 11) ist ein Karton, der mit zwei Linsen ausgestattet ist (Feyder und Rath-Wiggins, 2018). In Kombination mit einem Smartphone, das die Nutzer*innen in das vordere Ende der Brille einführen, bieten sie den kostengünstigsten Zugang zu einer Vielzahl von mobilen VR-Erlebnissen (Feyder und Rath-Wiggins, 2018).

4.2.3 Tracking und Bewegung

Als Tracking wird im Regelfall die Schätzung von Position und/oder Lage/Orientierung im Raum beschrieben (Dörner et al., 2019). Heutzutage kann die Positionsschätzung in der Regel mit hybriden 3DoF (Three-Degrees-of-Freedom) Sensoren vorgenommen werden (Dörner et al., 2019). Solche Sensoren sind heute in fast allen Smartphones, Tablets und VR-Brillen eingebaut (Dörner et al., 2019). 3DoF-Darstellungen beschreiben die drei in einer VR-Brille verfügbaren Drehwinkel (Magnor und Sorkine-Hornung, 2020). Dies ist unzureichend, um menschliche Bewegungen innerhalb des HMDs darzustellen, da Menschen sich auch entlang dreier Achsen bewegen können (Magnor und Sorkine-Hornung, 2020). Deshalb ist die 6DoF-Darstellung (vgl. Abb. 12) das Ziel der neuen

Generation von VR-Brillen, die eine realistische Reaktion auf menschliche Kopf- und Körperbewegungen ermöglicht und die Disparitätsprobleme des Stereo-360°-Formats löst, indem sie die Stereodarstellung in jeder Blickrichtung ermöglicht (Magnor und Sorkine-Hornung, 2020).

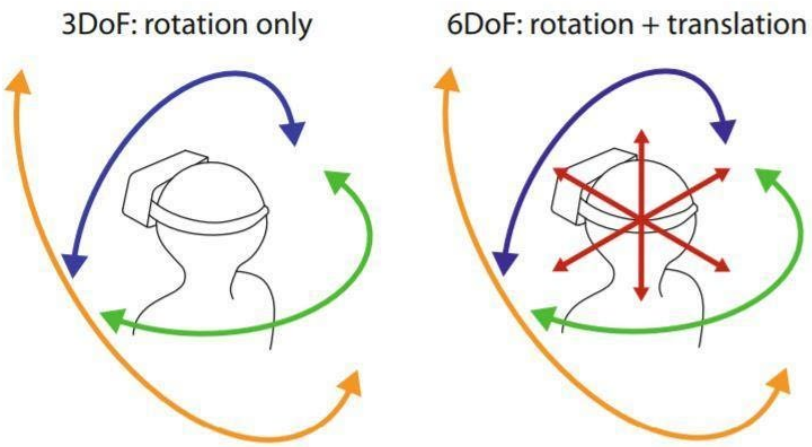


Abb. 12: Links: Mit dem „3DoF“ Tracking-System, kann sich der Benutzer von einem festen Punkt aus in der virtuellen Welt umschauen. Rechts: Mit dem „6DoF“ Tracking-System kann sich der Benutzer in der virtuellen Welt wie in der realen Welt bewegen (Magnor, Sorkine-Hornung, 2020)

4.2.4 Technische Eigenschaften der 360°-3D-Kameras

Zur Erstellung eines 360°-Videos ist ein spezielles Kamerasystem erforderlich, das aus mindestens zwei Kameralinsen besteht, die den Raum von einer Stelle, in alle Richtungen aufzeichnen (Windscheid und Gold, 2022). Bei zwei Linsen ist der Aufnahmewinkel immer größer als 180° (horizontal und vertikal), damit sich die Kamerabilder an ihren Rändern überlagern können (Windscheid und Gold, 2022). Bei Verwendung mehrerer Linsen, sind die Aufnahmewinkel dann entsprechend enger (Windscheid und Gold, 2022), (vgl. Abb. 13). 360°-3D-Kameras sind in der Regel mit mehreren Linsenpaaren ausgestattet, die jeweils etwa der Abstand der menschlichen Augen entspricht (Windscheid und Gold, 2022), (vgl. Abb. 14). Die perspektivische Verschiebung zwischen diesen Linsenpaaren, erzeugt eine visuelle räumliche Tiefe, die das menschliche Auge bzw. das Gehirn als dreidimensional wahrnimmt (Windscheid und Gold, 2022). Wenn alle Kamerabilder zu einem Video zusammengefügt und exportiert sind, liegt das fertige 360°-Video in einem sphärischen Projektionsformat vor (Windscheid und Gold, 2022). In der Regel wird das sogenannte Equirectangular

Panorama Format mit einem Seitenverhältnis von 2:1 verwendet (Windscheid und Gold, 2022). Beim Equirectangular Panorama Format wird das fertig zusammengestellte Bild mithilfe eines geeigneten 360°-Videoplayers auf die Innenfläche einer Kugel projiziert (Windscheid und Gold, 2022). Je nach Blickrichtung wird auf dem jeweiligen Monitor nur vom*von Betrachter*in anvisierten Bereich sichtbar (Windscheid und Gold, 2022).

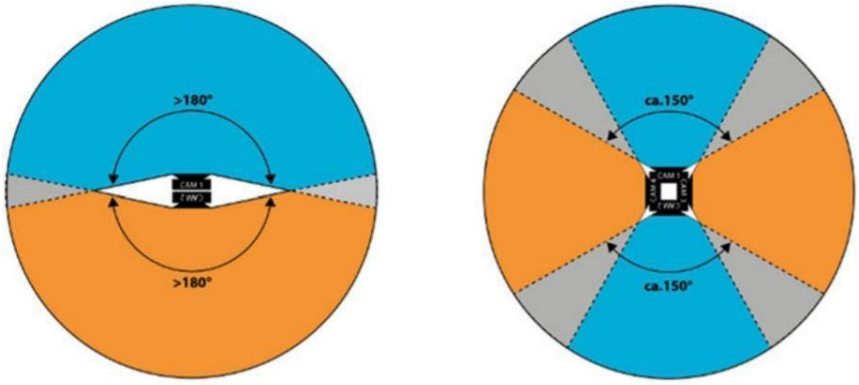


Abb. 13: Aufnahmewinkel einer 360°-Kamera. Links: zwei Kameralinsen, Rechts: vier Kameralinsen (Windscheid und Gold, 2022)

Abb. 14: Vuze Kamera ausgestattet mit 8 Kameralinsen (Amazon, 2022)

4.3 Visuelle Wahrnehmung (VR)

Ein wichtiges Merkmal des menschlichen Sehsystems ist das Sichtfeld. Es beträgt etwa 150° horizontal und 120° vertikal, wenn nur ein Auge verwendet wird, und vergrößert sich auf 180° horizontal und 120° vertikal, wenn beide Augen verwendet werden

(Burdea und Coiffet, 2003). Ein zentraler Teil dieses Sichtfeldes repräsentiert den Bereich der Stereopsis, indem das gleiche Bild von beiden Augen wahrgenommen wird (Burdea und Coiffet, 2003). Diese binokulare Überlappung beträgt horizontal etwa 120°(Burdea und Coiffet, 2003). Das Gehirn nutzt die horizontale Verschiebung der Bildposition, die von den beiden Augen registriert ist, um die Tiefe, oder die Entfernung zwischen dem Betrachter und dem virtuellen Objekt, das in der Szene dargestellt ist, zu messen (Burdea und Coiffet, 2003). Ein physiologisches Modell des menschlichen stereoskopischen Sehens ist in Abbildung 15 dargestellt (Burdea und Coiffet, 2003).

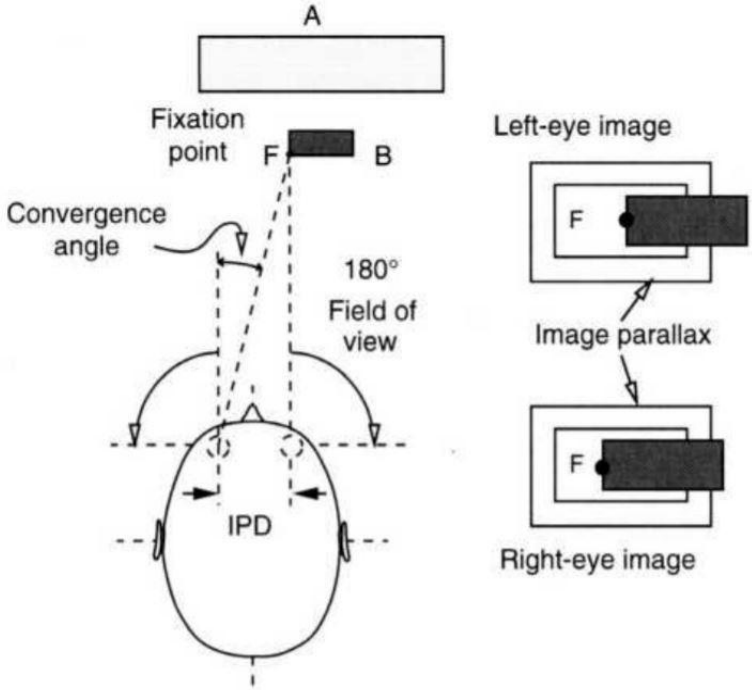


Abb. 15: Physiologisches Modell des stereoskopischen Sehens. IPD, Interpupillarer Abstand (Burdea und Coiffet, 2003)

Innerhalb des Sichtfelds registrieren die Augen die Objekte, die den Betrachter umgeben, wie das Objekt A, das hinter Objekt B liegt (Burdea und Coiffet, 2003). Die Augen konzentrieren dabei auf ein Merkmal von B, mit Fokus auf einen Fixpunkt F (Burdea und Coiffet, 2003). Der Winkel zwischen der Blickachse und der Linie zum Fixpunkt bestimmt den Konvergenzwinkel (Burdea und Coiffet, 2003). Dieser Winkel hängt vom Abstand zwischen den Pupillen des rechten und linken Auges, und wird als Interpupillardistanz (IPD) bezeichnet (Burdea und Coiffet, 2003). Er variiert bei

männlichen und weiblichen Erwachsenen in einem Bereich von 53-73 mm (Burdea und Coiffet, 2003). Der IPD ist die Grundlage, von der aus, eine Person die Entfernungen zu Objekten in der realen Welt interpretiert (Burdea und Coiffet, 2003). Je größer der IPD, desto größer ist der Konvergenzwinkel (Burdea und Coiffet, 2003). Der Punkt F erscheint horizontal verschoben zwischen dem rechten und dem linken Auge, da er sich im Verhältnis zu den beiden Augen in einer anderen Position befindet (Burdea und Coiffet, 2003). Diese Verschiebung wird als Parallax bezeichnet und muss von der VR-Grafik- und Stereosicht-Hardware nachgebildet werden, um das Gehirn bei der Interpretation der Tiefe in der simulierten Welt zu unterstützen (Burdea und Coiffet, 2003). Dazu müssen stereografische Bildschirme zwei leicht verschobene Bilder anzeigen (Burdea und Coiffet, 2003). Wenn zwei Bildschirme verwendet werden (wie z. B. bei HMDs), wird jeder Bildschirm sein Bild auf dem entsprechenden Auge anzeigen (Burdea und Coiffet, 2003). Bei Verwendung eines einzigen Bildschirms können die beiden Bilder getrennt dargestellt werden (bei autostereoskopischen Bildschirmen) (Burdea und Coiffet, 2003). Tiefenwahrnehmung, kombiniert mit einem breiten Sichtfeld, und ein hochauflösendes Bild sind wichtige Faktoren, die das subjektive Gefühl der Immersion bzw. Umhüllung in die simulierte Welt bestimmen (Burdea und Coiffet, 2003).

4.3.1 Präsenz und Immersion

Ein wesentliches Potenzial von VR liegt in der Möglichkeit, für den Nutzer*innen die Illusion der Anwesenheit in einer virtuellen Welt zu schaffen (Dörner et al., 2019). So soll dem*die Nutzer*in das Gefühl vermittelt werden, vollständig in die virtuelle Welt einzutauchen (Dörner et al., 2019). Präsenz ist der zentrale Begriff, der die psychischen Aspekte der VR-Erlebnisse beschreibt (Dörner et al., 2019). Er bezieht sich im weitesten Sinne auf das Gefühl, sich in der durch ein immersives VR-System dargestellten virtuellen Umgebung zu befinden (being there) (Dörner et al., 2019). Das allgemeine Gefühl der Präsenz umfasst drei verschiedene Teilaspekte (Dörner et al., 2019).

Erstens, die Ortsillusion beschreibt das Gefühl, sich vor Ort zu befinden (Dörner et al., 2019). Die Illusion, an einem Ort zu sein, hängt insbesondere von der Fähigkeit des immersiven VR-Systems ab, die Szene in Abhängigkeit vom Betrachter darzustellen (Dörner et al., 2019). Dreht der Nutzer zum Beispiel seinen Kopf um 90°- nach links, sollte

die virtuelle Umgebung zwar immer noch sichtbar sein, aber aus einer anderen Perspektive (Dörner et al., 2019).

Zweitens, die Plausibilitätsillusion tritt auf, wenn die Ereignisse in der simulierten Umgebung so wahrgenommen werden, als würden sie wirklich stattfinden (Dörner et al., 2019). Während die Standortillusion im Wesentlichen durch die Art der Darstellung verursacht wird, hängt die Plausibilitätsillusion stark vom Inhalt der simulierten Welt ab (Dörner et al., 2019). Die Plausibilitätsillusion bezieht sich insbesondere auf Ereignisse, die*den Benutzer*in betreffen, aber nicht von ihm initiiert wurden, z. B. ein plötzlich auf den*die Benutzer*in zufliegendes Projektil, das den*die Benutzer*in anspricht (Dörner et al., 2019).

Drittens, die Involviertheit bezeichnet den Grad der Aufmerksamkeit oder des Interesses der*des Benutzers*in an der simulierten Welt (Involvement wird, wie die Plausibilitätsillusion, hauptsächlich durch den Inhalt der virtuellen Umgebung hervorgerufen (Dörner et al., 2019). In einem immersiven VR-System kann eine*ein Benutzer*in beispielsweise das Gefühl haben, Teil der simulierten Welt zu sein (überzeugende Ortsillusion), und sich dennoch eher gelangweilt fühlen (niedrige Involviertheit) (Dörner et al., 2019).

Andererseits wird der Fachbegriff Immersion oder Umhüllung in einem technischen Sinn verwendet, sobald die Sinneseindrücke des VR-Teilnehmers durch ein oder mehrere Ausgabegeräte so umfassend wie möglich stimuliert werden (Dörner et al., 2019). Nach Slater und Wilbur (1997) gründet Immersion auf vier technischen Eigenschaften von Ausgabegeräten (Dörner et al., 2019):

a) die Sinneseindrücke des Menschen sollen möglichst ausschließlich durch den Computer generiert werden, d. h. der Nutzer soll weitestgehend von der realen Umgebung isoliert werden (Dörner et al., 2019).
b) möglichst viele Sinne sollen angesprochen werden (Dörner et al., 2019).
c) die Ausgabegeräte sollen den Nutzer vollständig umgeben, anstatt nur ein enges Sichtfeld zu bieten (Dörner et al., 2019).
d) zudem sollen die Ausgabegeräte eine „lebendige“ Darstellung bieten, z. B. durch hohe Auflösung und Qualität der Farbdarstellung (Dörner et al., 2019).

4.3.2 VR-Krankheiten

Unerwünschte Nebenwirkungen wie Kopfschmerzen, Übelkeit, Schwindel, Desorientierung usw. können bei der Nutzung von VR-Anwendungen entstehen (Dörner et al., 2019). Es ist bereits bekannt, dass eine lange Exposition gegenüber herkömmlichen Bildschirmen aufgrund einer Überanstrengung der Augen zu Kopfschmerzen führen kann (Dörner et al., 2019). Die Symptome sind bei der Nutzung von HMDs noch verstärkt, weil die Anzeige näher an den Augen ist, oder weil das stereoskopische Sehen noch eine Fusion erfordert (Dörner et al., 2019). Laut einer Studie von sind 10 Minuten HMD-Nutzung für das visuelle System genauso belastend wie 8 Stunden Sitzen vor einem Computerbildschirm (Dörner et al., 2019). Die Situation von Menschen mit Sehstörungen oder Problemen bei der Koordination der Augenmuskeln wird in der Folge meist noch verschlimmert, weshalb der Begriff (VR-Krankheiten) oder manchmal auch (Motionsickness) und (Cybersickness) verwendet wurde (Dörner et al., 2019). Auch wenn die Ursachen von VR-Krankheiten bislang nicht genau geklärt sind, konnten einige Faktoren identifiziert werden, die das Auftreten dieser Störungen begünstigen (Dörner et al., 2019).

Eine erste Gruppe von Faktoren ist abhängig vom Individuum, Alter, Geschlecht, aber auch individuelle Vorerfahrungen mit VR können das Auftreten der VR-Krankheit beeinflussen (Dörner et al., 2019). Eine zweite Gruppe von Einflussfaktoren bezieht sich auf den Bildkontrast und das damit verbundene Flackern, die Bildwiederholrate, Tracking-Fehler, die Qualität der Systemkalibrierung und die Verwendung stereoskopischer Bildschirme (Dörner et al., 2019). Je größer das Sichtfeld (und je mehr periphere Sicht möglich ist), desto häufiger wurde das Auftreten von VR-Krankheit beobachtet (Dörner et al., 2019). Wesentliche Faktoren sind zudem Latenzen, z. B. die zeitliche Verzögerung zwischen einer Kopfbewegung, dem Erkennen der neuen Kopfposition und der korrekten Darstellung des Bildes dieser neuen Kopfposition (Dörner et al., 2019). Um das Risiko des Auftretens von VR-bedingten Krankheiten zu reduzieren, müssen die Einflussfaktoren vermindert werden, z. B. durch Verkürzung der Latenzzeiten, durch die Verbesserung der technischen Eigenschaften der verwendeten VR-Geräte oder durch Aufklärung der Nutzer über mögliche Nebenwirkungen (Dörner

et al., 2019). VR-Krankheiten sind kein triviales Problem, sondern stellt eine echte Barriere für die Nutzung von VR dar, die bei der Entwicklung und Nutzung jedes VR-Produkts berücksichtigt werden muss (Dörner et al., 2019)

4.4 Anwendung der VR im Veranstaltungsbereich

Der virtuelle Raum bietet völlig neue Möglichkeiten für die Interaktion zwischen den Nutzern*innen und damit auch neue Gestaltungsmöglichkeiten für Inhalte, auf diese Weise wandelt sich der virtuelle Raum in einen vielseitigen Kreativitätsraum für die Nutzer*innen (Knoll, 2017). Es gibt zahlreiche Einsatzmöglichkeiten für VR im Veranstaltungsbereich, insbesondere auf Messen, wo VR-Systeme eingesetzt werden können, um komplexe oder noch nicht vorhandene Produkte zu präsentieren (Knoll, 2017). Produkte können auch in Situationen ausprobiert werden, die in der Realität nicht oder sehr teuer zu realisieren sind, z. B. kann jede*jeder Besucher*in eine Flugsimulation in einem virtuellen Cockpit erleben (Knoll, 2017).

Besondere Herausforderungen beim Einsatz von VR

Zum erfolgreichen Einsatz von Virtual Reality gibt es fünf Hinweise zu beachten (Knoll, 2017).

Hygiene Da HMD‘s direkt auf dem Kopf und im Gesicht des Benutzers getragen werden, sammeln sich Schweiß, Haare und Schmutz auf dem Gerät (Knoll, 2017). Linsen und Schaumstoffe sind dafür besonders anfällig (Knoll, 2017). Es wird daher empfohlen, die HMDs nach jedem Gebrauch zu desinfizieren (Knoll, 2017).

Platz. Das VR-Erlebnis sollte über einen ausreichenden Umgebungsraum verfügen, z. B. einen abgesperrten Bereich, in dem sich die Nutzer je nach Anwendung frei bewegen können (Knoll, 2017).

Sicherheit. Der Benutzer sollte die Möglichkeit haben, sich an das HMD zu gewöhnen und sich mit ihm vertraut zu machen (Knoll, 2017). Wichtige Faktoren sind hier ein sicherer Platz in der Nähe des Benutzers und Zeit, um sich an das Gerät anzuvertrauen (Knoll, 2017).

Betreuung. Da es noch keine etablierten Standards für den Umgang mit der Virtual Reality gibt, und da viele Nutzer VR zum ersten Mal die Geräte ausprobieren, sollte der

Anwender von einer Begleitperson betreut werden (Knoll, 2017). Diese Person sollte dem Nutzer das HMD an die Hand geben, bei Bedarf anpassen und Hinweise zur Nutzung liefern (Knoll, 2017). Dies gibt dem Nutzer zusätzliche Sicherheit, schützt die Technologie vor falscher Handhabung und sorgt für ein besseres Erlebnis (Knoll, 2017).

Dauer. Je nach Veranstaltung sollte das VR-Erlebnis zwischen zwei und zwanzig Minuten dauern (Knoll, 2017).

5 3D Audio

5.1 Auditive Wahrnehmung

Die wellenförmigen Wassermoleküle, die sich ausbreiten, wenn man einen Stein in ruhiges Wasser wirft, sind mit Schallwellen vergleichbar (Bühler et al., 2018). Diese Wellen sind schwingende Luftmoleküle, die das menschliche Ohr treffen und als Geräusche wahrgenommen werden (Bühler et al., 2018).

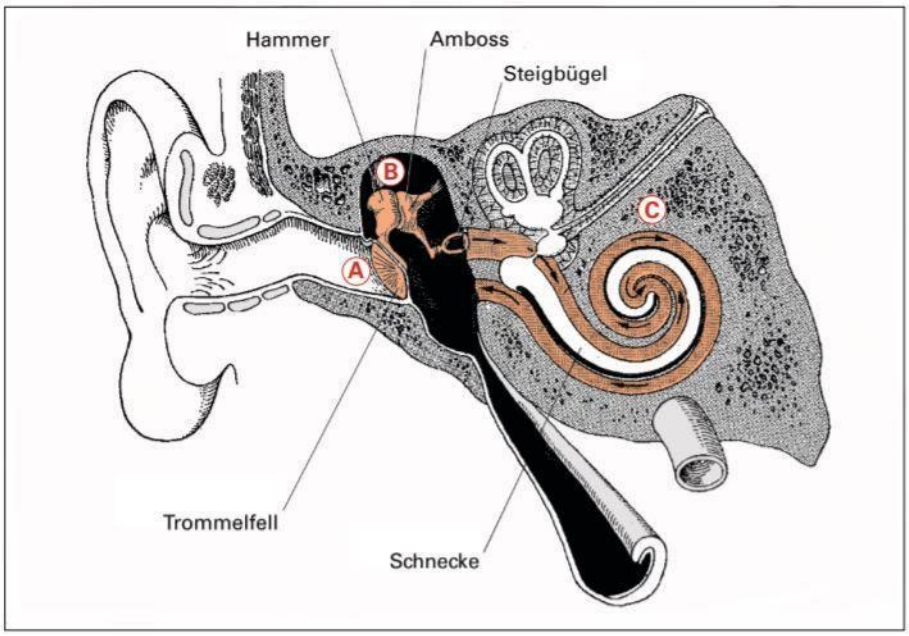


Abb. 16: Aufbau des menschlichen Ohrs (Bühler et al., 2018)

Das menschliche Ohr fungiert als Schallempfänger (Bühler et al., 2018), (vgl. Abb.16). Eintreffende Schallwellen gelangen durch den äußeren Gehörgang auf eine sensible dünne Membran, das Trommelfell A, das den Schalldruck in Schwingungen transformiert (Bühler et al., 2018). Die Schwingungen werden dann über drei Gehörknöchelchen, den Hammer, den Amboss und den Steigbügel, im Mittelohr B auf das Innenohr übertragen (Bühler et al., 2018). Die Schallübertragung von außen zum Innenohr führt zu einer Verstärkung des Schalldrucks um etwa den Faktor 60, dank der Eigenschaften der Gehörkomponenten des Innenohrs (Bühler et al., 2018). Der auditive Wahrnehmungsprozess befindet sich in der Schnecke im Innenohr C (Bühler et al., 2018). Sie enthält Härchen förmige Sinneszellen, die durch den Schalldruck verformt werden und diese Information über den Schneckennerv an das Gehirn weiterleiten

(Bühler et al., 2018). Dabei ermittelt das Gehirn die Lautstärke, die Tonhöhe und die Schallrichtung (Bühler et al., 2018).

Immersive Audio Wahrnehmung

Immersion bezieht sich akustisch auf Geräusche, die aus allen Richtungen rund um die*den Zuhörer*in eintreffen (Roginska und Geluso, 2018). Hörbare Schallquellen breiten sich aus und werden von Oberflächen um die*den Hörer*in herum reflektiert (Roginska und Geluso, 2018). Die allgemeine Bedeutung von Immersion in Audio und Akustik bezieht sich jedoch auf das psychologische Gefühl, von bestimmten Schallquellen sowie von Umgebungsgeräuschen umgeben zu sein (Roginska und Geluso, 2018). Räumliche Audiotechniken, insbesondere 3D-Audio, können ein immersives Erlebnis vermitteln, da virtuelle Schallquellen und Schallreflexionen überall im Raum um die*den Hörer*in herum erscheinen können (Roginska und Geluso, 2018). Die wahrgenommene Entfernung einer Schallquelle wird hauptsächlich durch die akustischen Eigenschaften von Schallpegel und Nachhall bestimmt (Roginska und Geluso, 2018). Normalerweise erreichen Schallquellen die*den Hörer*in nicht nur durch eine ungehinderte Wellenfront (Direktschall), sondern auch durch Reflexion und Beugung an nahen gelegenen Objekten wie dem Boden, Wänden oder anderen Oberflächen (indirekter Schall oder Nachhall) (Roginska und Geluso, 2018).

5.2 3D Audio

In den letzten Jahrzehnten hat sich die räumliche Wiedergabe von Tonsignalen von einfachem kanalbasierten 2-Kanal-Stereo zu Surround-Sound mit 5.1 bzw. 5 Lautsprecher und ein Subwoofer, Lautsprecher und in den letzten Jahren zu 3D-Sound Scene-basiert mit z. B. (22.2) entwickelt (Schoeffler et al., 2015). Denn mit zunehmender Anzahl an Lautsprechern wird eine Verbesserung der räumlichen Wahrnehmung und der Einhüllung der*des Zuhörers*in erwartet (Schoeffler et al., 2015). Authentische Reproduktion von Raumklang (inklusive Musik) ist ein langjähriges Forschungsthema in der Audiowelt (Schoeffler et al., 2015). Von Anfang an bestand das Ziel darin, die wesentlichen Merkmale eines Raumklangs in einer Aufnahme zu erfassen und sie an einem anderen Ort und zu einer anderen Zeit so wiederzugeben, dass die*der Hörer*in

einen ähnlichen akustischen Eindruck wie am ursprünglichen Ort und zur ursprünglichen Zeit empfindet (Schoeffler et al., 2015). Ein wichtiger Meilenstein auf dem Weg zu diesem Ziel war die Einführung des Zweikanal-Stereosystems durch Blumlein im Jahr 1931 (Schoeffler et al., 2015). Nach mehreren Jahrzehnten aktiver Forschung wurden signifikante Fortschritte mit Techniken erreicht, bei denen eine erweiterte Anzahl an Lautsprechern, für die Wiedergabe verwendet wurden, was zu einem stark verbesserten Gefühl führte, vom Sound umgeben zu sein „Umhüllung“ (Schoeffler et al., 2015).

5.2.1 Ambisonics

Ambisonics wurde in den frühen 70er Jahren von Michael Gerzonau entwickelt und ist eine 3D-Audio-Spatialization-Technologie, welches ermöglicht, 3D-Soundfelder auf eine flexible Weise zu rendern (Moreau et al., 2006). Das Ambisonic-Surround-Sound-System ist im Wesentlichen eine zweiteilige technologische Lösung für das Problem der Kodierung von Schallrichtungen (und -amplituden) und deren Wiedergabe über praktische Lautsprechersysteme auf eine Weise, die die*dem Zuhörer*in das Gefühl vermittelt, den Original-Ton an der richtigen Stelle zu hören (Malham, 1990). Die Beschallung kann über eine 360°-Horizontalfläche (pantophonic systems) oder über die gesamte Sphäre (periphonic systems) erfolgen (Malham, 1990). Systeme, die zur Übertragung der aufgezeichneten Informationen sogenannte (B-Format) Signale, benötigen drei bzw. vier Kanäle für die horizontale bzw. vollsphärische Kodierung der Töne (Malham, 1990). Für die Wiedergabe sind vier oder mehr Lautsprecher erforderlich, je nachdem, ob es sich um eine pantophone oder periphone Wiedergabe handelt (Malham, 1990). Das praktische Minimum ist vier für die horizontale Beschallung und sechs, wenn die Höhe noch benötigt wird (Malham, 1990). Während kanalbasierten, diskreten Surround-Sound-Verfahren durch ihre Aufnahmeanordnung auf ein bestimmtes Wiedergabesystem beschränkt sind, bietet das Ambisonic-Format mehr Flexibilität (Malham, 1990). Sobald die Aufnahmen im Ambisonic-Format kodiert sind, lassen sie sich binaural über Kopfhörer oder über jedes Lautsprechersystem beliebige wiedergeben werden (Hoffbauer, 2021).

5.2.2 Tetraedermikrofone

Die Kombination von 4 Tetraederförmig angeordneten Kapseln, auf der das erster Ordnung Mikrofon (engl. First Order Ambisonics, kurz FOA Mikrofon) basiert, wurde von Michael Gerzon im Jahr 1973 entwickelt (Pfanzagl-Cardone, 2020). Die Idee dahinter ist, dass die Kapseln in möglichst direkter Nähe zueinander liegen (Pfanzagl-Cardone, 2020). Idealerweise am gleichen Punkt des Soundfeldes, was aber physikalisch nicht möglich ist (Pfanzagl-Cardone, 2020). Der Tetraeder (vgl. Abb.17) repräsentiert die bestmögliche Lösung für dieses Problem (Pfanzagl-Cardone, 2020).

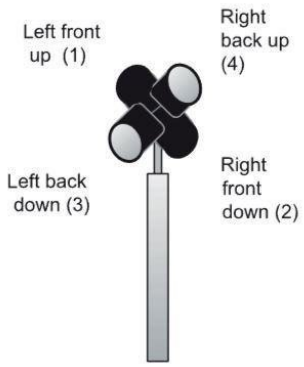


Abb. 17: Beispiel ein tetraedrisches Mikrofonsystem (Paterson und Lee, 2021)

Das Mikrofon liefert vier Richtungssignale (vgl. Abb.18), die Kapseln eins (LFU) und vier (RBU) sind um 45° nach oben ausgerichtet, die Kapseln zwei (RFD) und drei (LBD) sind um 45° nach unten ausgerichtet (Paterson und Lee, 2021). Die vier diskreten Ausgänge eines tetraedrischen Mikrofons bilden die vier Komponenten des Ambisonic A-Formats (Paterson und Lee, 2021).

Channel	Orientation	Abbreviation
1	Left front up	LFU
2	Right front down	RFD
3	Left back down	LBD
4	Right back up	RBU

Abb. 18: Tetraedrische Nummerierung der Mikrofonausgangskanäle (Paterson und Lee, 2021)

Durch Matrixbildung können diese 4 sogenannten A-Format-Signale im B-Format umgewandelt werden (Paterson und Lee, 2021). Aus diesen 4 B-Format-Signalen lassen sich Richtungsmuster erster Ordnung in jeder gewünschten Richtung abgeleitet werden, wenn die entsprechende Verarbeitung durchgeführt wird (Paterson und Lee, 2021).

5.2.3 Ambisonics erster Ordnung

Ambisonics erster Ordnung wird heute durch das Internet stark vorangetrieben, unterstützt durch YouTube, Facebook, 360°-Audio-Videoaufnahmen und -wiedergaben sowie VR-Spiele (Zotter und Frank, 2019). Diese Renaissance liegt in den Vorteilen der kompakten Hauptmikrofon-arrays (vgl. Abb.19, 20, 21), die den gesamten Umgebungsschall in nur vier Audiokanälen aufzeichnen können (z. B. Zoom H3-VR, Sennheiser AMBEO VR Mic, Røde NT-SF1) (Zotter und Frank, 2019), und ermöglichen, eine einfache Rotierung der Soundszene, sodass Surround-Audioszenen gerendert werden können, z. B. auf Kopfhörern mit Headtracking, kopfgetragenen AR/VR-Geräten oder mobilen Geräten (Zotter und Frank, 2019).

Abb. 19: Zoom H3-VR

(Zoom -Webseite)

Abb. 20: Sennheiser AMBEO VR Mic

(Sennheiser -Webseite)

Abb. 21: Røde NT-SF1

(Røde Webseite)

5.3 3D Aufnahme

Um eine 3D-Aufnahme zu realisieren, muss mindestens ein tetraedrisch angeordnetes Mikrofon verwendet werden, die eine sogenannte Ambisonic-Aufnahme im A-Format aufzeichnet (Paterson und Lee, 2021). Diese Aufnahme kann in vier Komponenten verarbeitet - kodiert - werden: eine kugelförmige omnidirektionale Komponente, die die notwendige null Ordnung (W) darstellt, und drei bidirektionale Komponenten (X, Y und Z), die die erste Ordnung definieren (Paterson und Lee, 2021), (vgl. Abb. 22).

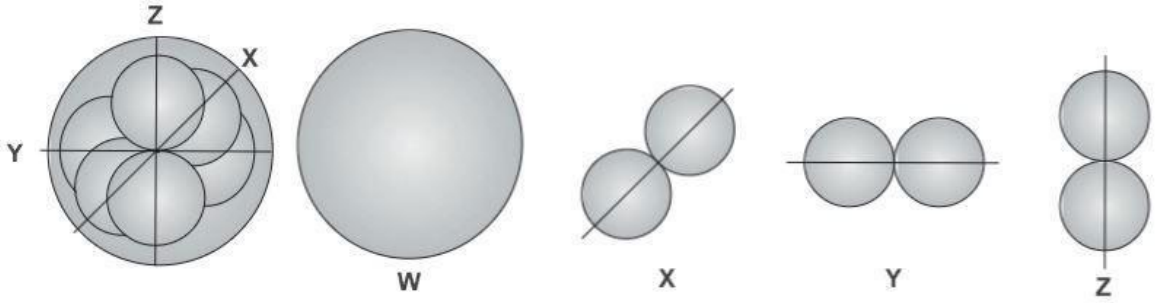


Abb. 22: Ambisonic B-Format-Komponenten; übereinandergelegt (ganz links) und separat (Mitte bis rechts) W, X, Y und Z dargestellt (Paterson und Lee, 2021)

Diese vier Komponenten mit den Bezeichnungen W, X, Y und Z werden als sphärische Harmonische bezeichnet und können zusammen (siehe Abb. 23) das sogenannte Ambisonic-B-Format erster Ordnung bilden (Paterson und Lee, 2021). Die maximale Anzahl der räumlichen Komponenten, die ein Soundfeld definieren, sind proportional zur Anzahl der Systemkanäle und bestimmt auch zur „Ordnungszahl" des Systems (Paterson und Lee, 2021). Die Anzahl der Ambisonic-Kanäle hängt von der verwendeten Ordnung der Ambisonics ab (Paterson, Lee, 2021). Höhere Ordnungen erfordern eine größere Anzahl von Kanälen und bieten in der Regel eine höhere räumliche Auflösung auf Kosten der Datenspeicherung, der Rechenleistung und erhöhter Komplexität (Paterson, Lee, 2021).

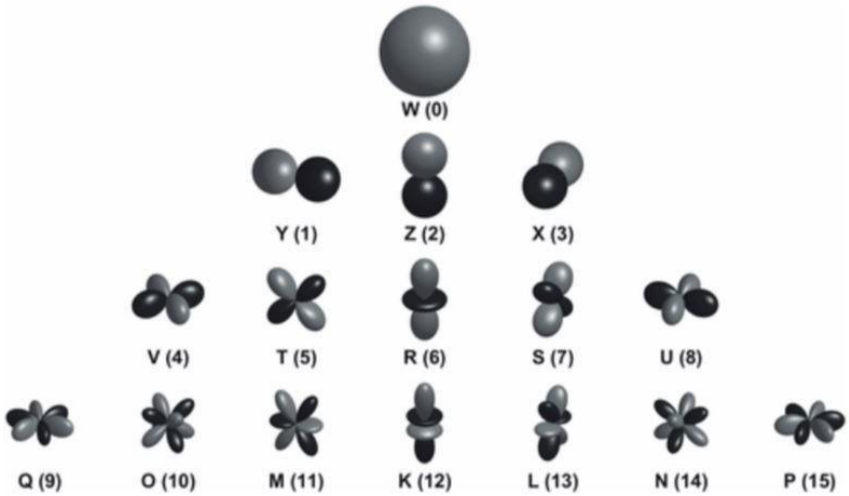


Abb. 23: Die SN3D normalisierten sphärischen harmonischen Funktionen, welche in Ambisonics bis zur dritten Ordnung verwendet werden, gekennzeichnet mit alphabetischer (FuMa) und ACN-Notation (Paterson, Lee, 2021)

Die obere Abbildung zeigt die ersten drei Ambisonics-Ordnungen, beginnend mit der 0. Ordnung (Brühwiler, 2022). Aus der Kombination der 0. Ordnung und der zweiten Reihe ergibt sich die 1. Ordnung mit insgesamt vier Kanälen (Brühwiler, 2022). Wenn die dritte Reihe zu dieser Ordnung hinzugefügt wird, ergibt sich die 2. Ordnung mit 9 Kanälen und schließlich die vollständige, die 3. Ordnung mit 16 Kanälen, die sich aus der Kombination aller vier Reihen ergibt (vgl. Abb. 23) (Brühwiler, 2022).

5.3.1 B-Format Konvertierung

Wie bereits erwähnt, kann das B-Format für herkömmliche Stereo- und Surround-Anwendungen problemlos vom A-Format abgeleitet werden (Paterson und Lee, 2021). Das B-Format wird von kommerziellen Surround-Decodern, Immersive-Sound-Software sowie von Streaming-Diensten und Übertragungsformaten wie MPEG-H unterstützt (Paterson und Lee, 2021). Die Kodierung vom A-Format in das B-Format erfolgt mithilfe der folgenden Summationsmatrix (hier ungewichtet dargestellt):

W = LF + RF + LB + RB (Summe aller Signale, omnidirektional)

X = LF – LB + RF – RB (Vorne/Hinten, Tiefeninformationen)

Y = LF + LB – RF – RB (Links/Recht, horizontal)

Z = LF – LB – RF + RB (Oben/Unten, vertikal)

(Paterson und Lee, 2021).

Mikrofonanbieter wie Sennheiser und RØDE stellen für die Matrizierung ihrer A-Format Mikrofonsignale jeweils ein eigenes Plugin zur Verfügung (Brühwiler, 2022). Mithilfe der Plugins können die Signale ins B-Format konvertiert werden (Brühwiler, 2022).

5.3.2 Output-Formate

Die beiden Ausgabestandards des Ambisonics-B-Formats sind das AmbiX-Format, was für „Ambisonics exchangeable“ steht und das Furse-Malham (kurz FuMa) (Brühwiler, 2022). Sie sind zwar ähnlich, aber nicht austauschbar: Sie unterscheiden sich in der Reihenfolge, in der die vier Kanäle angeordnet sind (Waves, 2017). Die Kanalanordnung des AmbiX-Formats wird mit ACN (engl. ambisonic channel order) abgekürzt und ordnet

die vier Kanäle in der Reihenfolge W-Y-Z-X, während FuMa die Reihenfolge W-X-Y-Z verwendet (Brühwiler, 2022). Die Pegelunterschiede zwischen den beiden Formaten sind auch auf die unterschiedliche Normalisierung der Lautstärke angewiesen (Brühwiler, 2022). AmbiX verwendet die Semi-Normalisierung SN3D, die sicherstellt, dass der erste Kanal W bei der Codierung der vier Kanäle von keinem anderen Kanal im Pegel überschritten wird (Brühwiler, 2022).

5.3.3 Binaural Audio über Kopfhörer

Binaurale Signale bestehen aus zwei Kanälen (Magnor und Hornung, 2020). Im Gegensatz zur Stereowiedergabe über Kopfhörer enthalten binaurale Signale jedoch integrierte räumliche Informationen in Form von Zeit, Intensität und spektraler Färbung Elemente, die die natürliche menschliche Hörwahrnehmung simulieren und optimieren (Magnor und Hornung, 2020). Für die Wiedergabe binauraler Signale sind Kopfhörer oder Lautsprecher vorgesehen (Magnor und Hornung, 2020). Wenn binaurale Signale korrekt aufgenommen, synthetisiert und wiedergegeben sind, vermitteln sie einen starken Eindruck von räumlichen Klängen, als ob sie in einer natürlichen Hörumgebung auftreten würden (Magnor und Hornung, 2020).

HRTFs

Die kopfbezogenen Übertragungsfunktionen (engl. Head-related transfer functions kurz: HRTFs) sind eine entscheidende Komponente des binauralen Klangs und bestimmen zu einem großen Teil die Qualität jedes binauralen 3D-Klangs, der wahrgenommen ist (Paterson und Lee, 2021). Die menschlichen Ohrmuscheln sind so einzigartig wie die Fingerabdrücke, und daher sind auch die HRTFs höchst unterschiedlich (Paterson und Lee, 2021). Modernen räumlichen Audiowiedergabesysteme arbeiten meist mit generische HRTFs, die nicht unbedingt das beste räumliche Klangerlebnis bieten (Paterson und Lee, 2021). Die Verwendung generischer HRTFs führt häufig zu Verwirrungen in Bezug auf die Vorderseite und die Rückseite sowie zu einer Lokalisierung im Kopf, die die Wahrhaftigkeit des räumlichen Bildes beeinträchtigt (Paterson und Lee, 2021). Die Anwendung personalisierter HRTFs hilft, einige dieser Probleme zu beseitigen, ist aber ein extrem mühsamer Aufnahmeprozess (Paterson und Lee, 2021). Eine HRTF enthält sowohl die Magnituden-

als auch die Phaseninformation der resultierenden Schallwelle nach ihrer komplexen Interaktion mit dem Kopf, dem Oberkörper, der Schulter und den Ohrmuscheln (Paterson und Lee, 2021).

Bewegung und Dynamic

In einer virtuellen Hörumgebung kann die Bewegung das Ergebnis einer Kombination aus einem der beiden folgenden Elemente sein bei der binauralen Synthese: Bewegung des Hörers und Bewegung der Schallquelle (Roginska und Geluso, 2018). Durch Drehen des Kopfes kann ein Hörer seine Ausrichtung in Bezug auf Drehen, Nicken und Wanken ändern (Roginska und Geluso, 2018). Die Änderung der Kopfausrichtung beeinflusst die relative Position der Schallquelle in Bezug auf den Kopf der*des Zuhörers*in. Zum Beispiel, eine Änderung der Drehung des Kopfs führt zu einer Änderung des relativen Azimuts der wahrgenommenen Schallquelle (Roginska und Geluso, 2018).

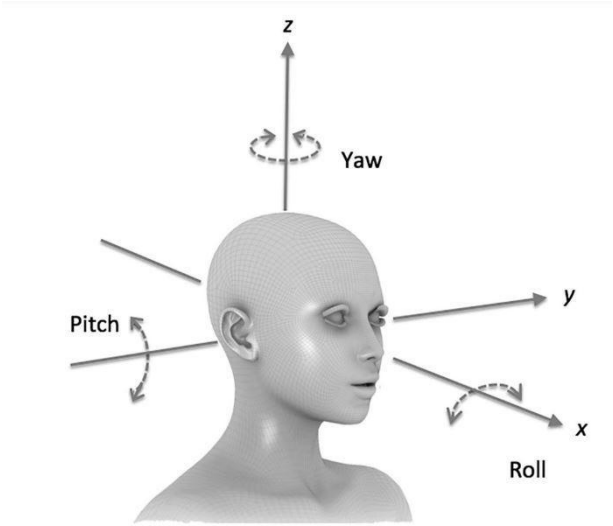


Abb. 24: Dreidimensionale Kopfrotation: Drehen (engl. yaw), Nicken (engl. pitch) und Schwenken (engl. roll) (Roginska und Geluso, 2018)

Es gibt viele Techniken und Technologien, um die Position eines Hörers zu ermitteln (Roginska und Geluso, 2018). Headtracker messen die Position eines Zuhörers relativ zu einem Referenzpunkt. 3DOF-Tracker (Three degrees-offreedom) messen nur die Ausrichtung des Kopfes (vgl. Abb. 24) „Drehen“, „Wanken“, „Nicken“, (Roginska und Geluso, 2018). Während sechs Freiheitsgrade (6DOF) messen sowohl die Ausrichtung als auch die x, y und z Position des Zuhörers (Roginska und Geluso, 2018).

Binaural Decoder

Der binaurale Decoder berechnet virtuelle Lautsprechersignale als lineare Kombination der Ambisonics-Signale gemäß einer vorgegebenen Dekodiermatrix (Kronlachner, Zotter, 2013). Diese virtuellen Lautsprechersignale werden mit ihren individuellen Stereo-Impulsantworten gefaltet und modellieren so den Übertragungsweg von der Lautsprecherposition zum linken und rechten Ohr der*des Zuhörers (Kronlachner, Zotter, 2013).

5.4 Plug-ins zur 3D Audioverarbeitung

Plug-ins sind spezielle Softwareerweiterungen, die in Audioproduktionsprogramme (engl. Digital Audio Workstations, kurz DAWs) über standardisierte Schnittstellen wie (VST, engl. Virtual Studio Technology) installiert und verwendet werden können (Weinzierl, 2008).

Einige Bespiele von verwendeten Plug-ins zur 3D-Audioverarbeitung sind hier zitiert:

5.4.1 IEM Plug-in Suite

Eine weitere Open Source Plugin Suite zur ambisonsichen 3D-Audio signalverarbeitet bietet das Institut für Elektronische Musik und Akustik in Graz an (IEM 2021). Alle Ambisonics Plugins aus dieser Suite können bis zur 7. Ordnung verwendet werden und auf der IEM Website mit Beschreibungen gefunden werden (IEM, 2021).

5.4.2 Facebook Spatial Workstation

Die Facebook 360 Spatial Workstation ist eine komplette Suite, die es Sounddesignern ermöglicht, Audioquellen einzubinden, zu schwenken und mit Szenenelementen zu synchronisieren und in eine einzige Ambisonic-Datei zu rendern, die auf Facebook, Oculus Video und anderen Plattformen wiedergegeben werden kann (Thakur, 2017). Ursprünglich von Two Big Ears entwickelt, ist Facebook 360 Spatial Workstation jetzt ein kostenloses Tool, das vom Audio 360-Team bei Facebook bereitgestellt wird (Thakur, 2017). Leider hat die Firma Meta die Unterstützung für die Facebook Spatial Workstation Suite am 16. Mai 2022 beendet (Github, 2022).

5.5 Anwendung der 3D Audio im Veranstaltungsbereich

Für das Publikum bieten diese neuen Systeme eine höhere Lokalisierungsgenauigkeit und ermöglichen eine verbesserte audiovisuelle Einheitlichkeit in einem sehr großen Teil des Veranstaltungsbereichs (Paterson und Lee, 2021). Die Fähigkeit, Instrumente im Raum zu differenzieren - räumliche Trennung - ist ebenfalls stark verbessert (Paterson und Lee, 2021). Für Live-Tontechniker, für die Links/Rechts-Systeme häufiger Dual-Mono-Systeme sind, öffnen diese neuen Technologien die Türen zu neuen kreativen Möglichkeiten (Paterson und Lee, 2021).

Um einen erfolgreichen Einsatz dieser Technologien in einem Live-Kontext zu gewährleisten, ist nicht nur eine einfache Umstellung von kanalbasiertem auf objektbasiertes Audio erforderlich, sondern ein umfassender Systemansatz (Paterson und Lee, 2021).

Dies erfordert:

- Räumliche Audioalgorithmen, die an die besonderen Bedingungen von Live-Sound angepasst sind (Paterson und Lee, 2021).
- eine Methodik für den Entwurf von Lautsprechersystemen und Bewertung, mit denen das für einen bestimmten Veranstaltungsort erforderliche Lautsprechersystem bestimmt und optimiert werden kann (Paterson und Lee, 2021).
- Angepasste Mischtechniken und kreative Werkzeuge, wie z. B. 3D-Nachhall, die an große Wiedergabesysteme und Vorrangbedingungen angepasst sind (Paterson und Lee, 2021).
- Signalprozessoren und Software, die in der Lage sind, eine große Anzahl von Ein- und Ausgängen mit einer geringen Latenzzeit von in der Regel unter 5 ms zu verarbeiten (Paterson und Lee, 2021).

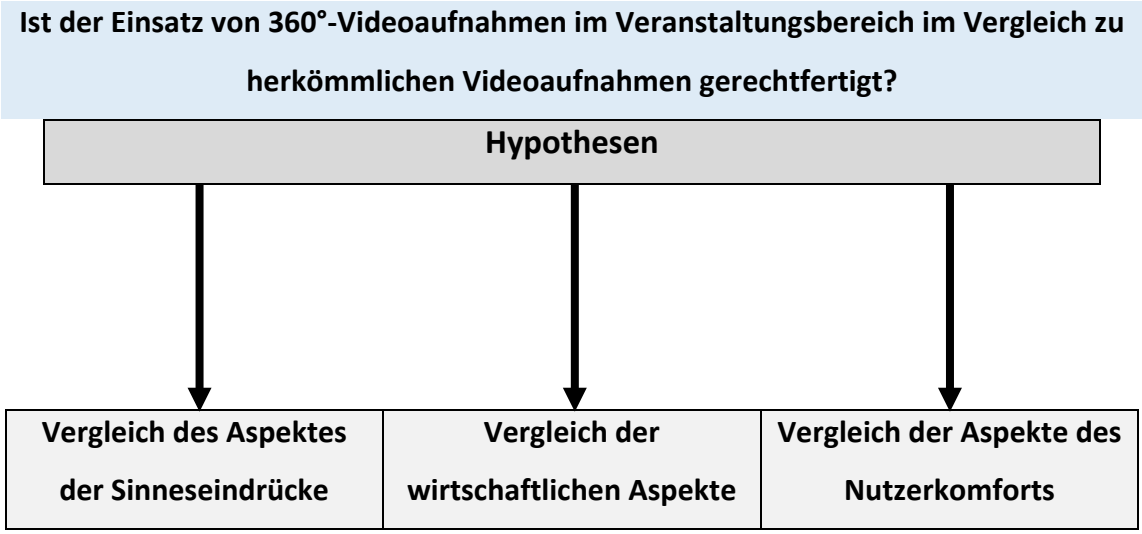
Zahlreiche Lösungen sind derzeit auf dem Markt zu finden, entweder von Lautsprecherherstellern oder von unabhängigen Herstellern oder Softwareherstellern (Paterson und Lee, 2021). Alle diese Lösungen verwenden einen objektbasierten Lösungsansatz, unterscheiden sich jedoch durch den verfügbaren Panning-Algorithmus,

die 3D-Wiedergabe, das Design des Lautsprechersystems sowie die Software- und Hardwareimplementierung (Paterson und Lee, 2021).

6 Methodik

Im ersten Teil dieses Kapitels wird die Forschungsfrage dargestellt. Aus dieser werden Hypothesen abgeleitet, deren Gültigkeit im Laufe der Arbeit durch ein empirisches Vorgehen überprüft werden sollen. Zunächst wird das methodische Vorgehen und die empirische Untersuchung beschrieben. Schließlich folgt eine Begründung für die Wahl der Methode sowie Informationen zum Aufbau des Fragebogens mit zugehörigen Erläuterungen. Zur Durchführung eines umfassenden Vergleichs zwischen der konventionellen Aufnahmemethode in 2D und der Aufnahme mit einer 360°-Kamera wurde eine Online-Umfrage entwickelt. Anhand dieser sollte ermittelt werden, inwieweit sich diese neue Technologie auf den Betrachter*innen auswirkt und ob es sinnvoll ist, diese neue Art der Aufnahme im Veranstaltungsbereich einzusetzen. Um diese Fragen zu beantworten, werden im Folgenden der Forschungsgegenstand und die Vorgehensweise erläutert.

6.1 Forschungsfrage



6.1.1 Vorgehensweise und Vorüberlegungen

Im Vorfeld dieser Forschungsarbeit musste zunächst geklärt werden, ob eine qualitative oder quantitative Studie als Forschungsmethode besser geeignet ist. Bei der Beantwortung der Forschungsfragen sollte bereits vor der Durchführung der Umfrage

festgelegt werden können, welche genauen Aspekte der Untersuchung in dieser Studie untersucht werden sollen.

Demzufolge müssen einzelnen die Aspekte jedes des zu untersuchenden Bereiches genau kategorisiert werden, um daraus realistische, aussagekräftige, allgemeine und objektive Schlussfolgerungen ziehen zu können. Daher fiel die Wahl auf eine quantitative Auswertungsmethode. Die Umfrage beinhaltete jedoch eine halboffene und eine offene Frage, die qualitativ ausgewertet werden.

Quantitative Forschung verfolgt das Ziel, das Verhalten in Form von Modellen und numerischen Daten so genau wie möglich zu beschreiben, um eine Art von Verhalten vorhersehbar zu machen. Daraus folgt, dass diese Methode das menschliche Verhalten anhand der untersuchten Fakten beobachtet, um daraus Schlussfolgerungen zu ziehen. Gewonnene Erkenntnisse werden dann auf nicht untersuchte Fälle angewendet. So kann die Richtigkeit der zuvor vorgestellten Hypothesen und Theorien überprüft werden. Die qualitative Forschung dient dazu, durch offene Fragen, Aspekte zu identifizieren, die bei der Konstruktion des Fragebogens nicht berücksichtigt wurden. Auf diese Weise lassen sich neue Erkenntnisse und Informationen generieren (Hollenberg, 2016).

Für die Studie wurden 15 Probanden (aus dem Studien- und Bekanntschaftskreis) eingeladen, die Videos anzuschauen. Im Anschluss sollten sie mit dem Beantworten der Fragebögen ihre Meinung zu den unterschiedlichen Videoformaten geben.

Die untersuchten Kriterien gliedern sich in vier Gruppen: Ein allgemeiner Teil umfasst demografische Fragen, Fragen zum verwendeten Equipment während der Untersuchung, sowie eine allgemeine Frage zur Erfahrung mit VR. Im zweiten Teil sollen die Probanden Aspekte der Sinneseindrücke bewerten. Der dritte Teil enthält Fragen zur Bewertung des wirtschaftlichen Aspekts und schließlich Fragen zur Bewertung der Ergonomie bzw. des Komfortaspekts bei der Verwendung einer VR-Brille.

Bei der Bewertung des Aspekts der Sinneseindrücke liegt der Schwerpunkt auf der visuellen und auditiven Wahrnehmung. Für die visuelle Wahrnehmung sollen die Probanden die Auflösung (Bildqualität) und die entstehenden Stimuli und Gefühle während der Betrachtung vergleichen. Bei den Gefühlen wird zwischen dem Präsenzgefühl, d.h. der Illusion, in einer virtuellen Welt vor Ort zu sein, und dem

Umhüllungsgefühl bzw. Immersion, d.h. Teil der simulierten Welt zu sein, unterschieden. Solche psychovisuellen Aspekte werden durch Befragung untersucht und anschließend mit den Stimuli der üblichen 2D-Aufnahme verglichen. Bei der auditiven Wahrnehmung wird zwischen 3D-Audioaufnahme und Stereoaufnahme verglichen. Der Unterschied besteht darin, dass sich der Pegel bei der 3D-Audioaufnahme mit der Kopfbewegung ändert und der Klang dabei natürlicher wirken soll, während der Stereoton statisch bleibt und damit weniger authentisch klingen soll. In der Studie wird dann untersucht, ob diese Erfahrung einen besseren Höreindruck vermittelt.

Der wirtschaftliche Aspekt ist ein wichtiger Bestandteil der Befragung. Für ein herkömmliches 2D-Video fallen fast keine Investitionen an, im Gegensatz zu einem VR-Erlebnis. Die Intensität dieser VR-Erlebnisse hängt von der Gestaltung der Inhalte (vgl. Abb.26), der Produktionstechnik und -qualität sowie, auf der Verbraucherseite, von der Art der verwendeten VR-Brille und der Leistung der verwendeten Endgeräte ab (Feyder und Rath-Wiggins, 2018).

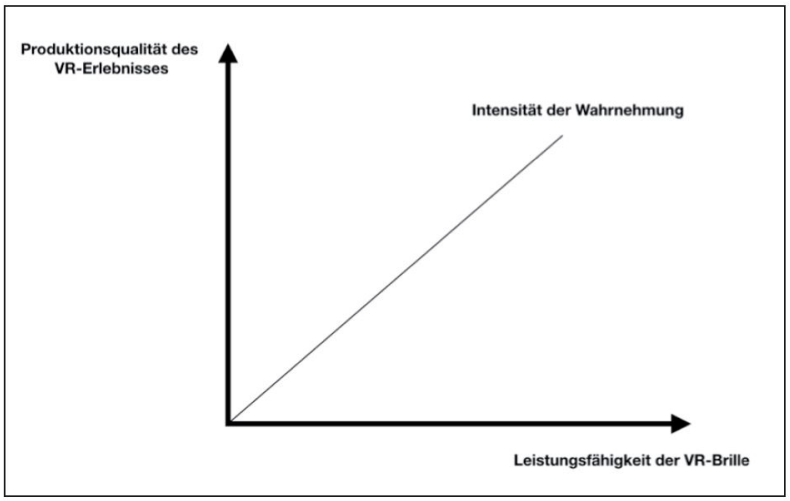


Abb. 25: Intensität der VR-Wahrnehmung (Feyder und Rath-Wiggins, 2018)

Ein Unterschied besteht zwischen der Verwendung einer Mobile VR-Brillen wie das Cardboard oder Samsung Gear zum Beispiel, welches ein kostengünstiges VR-Erlebnis bietet. Der Kauf einer stationären VR-Brille mit integrierter Anzeige, die über ein größeres mit höherer Auflösung verfügt sowie Tracking Sensoren zur Bewegungsverfolgung. In diesem Fall muss die*der Verbraucher*in derzeit ziemlich viel

Geld investieren, außerdem braucht er einen leistungsfähigen Rechner oder Spielekonsolen (Feyder und Rath-Wiggins, 2018). In der Studie wird daher gefragt, ob die Teilnehmer bereit sind, Geld für diese Technologie auszugeben. So lässt sich abschätzen, ob von Seiten der Betrachter*in ein Interesse besteht, VR-Brillen im Alltag zu verwenden.

Bei der Bewertung des Komforts bzw. des ergonomischen Aspekts der Nutzung von VR-Brillen spielt das Tragen eines solchen Geräts eine große Rolle, da die Menschen im Allgemeinen nicht daran gewöhnt sind, ein derart großes Gerät auf dem Kopf zu tragen. Darüber hinaus können bei der Nutzung von VR-Anwendungen unerwünschte Nebenwirkungen wie Kopfschmerzen, Übelkeit und Schwindel auftreten, die meist von Alter und Geschlecht abhängig sind. Dies stellt ein reales Hindernis für die Nutzung von VR dar. Deshalb wurden bei der Erstellung der Umfrage demografische Fragen gestellt, um zu ermitteln, ob die so genannten VR-Krankheiten in einer bestimmten Altersgruppe oder bei einem bestimmten Geschlecht stärker ausgeprägt sind.

Diese Untersuchung soll Gemeinsamkeiten und Unterschiede aufzeigen und verdeutlichen. Diese können hilfreiche Hinweise und damit Schlussfolgerungen über notwendige Maßnahmen und Ursachen liefern.

Zur Datenerhebung wird ein Fragebogen verwendet. In diesem sind 11 geschlossene Fragen, 2 halboffene Fragen und eine offene Frage enthalten. Die benötigte Antwortzeit beträgt etwa 10 Minuten. Für den direkten Vergleichsaspekt wurde eine fünfstufige Skala gewählt, da sie je nach Meinung des Befragten ein breites Spektrum an Antwortmöglichkeiten abdeckt, mit zwei Stufen für Nichtgefallen oder Zufriedenheit und einem neutralen Mittelwert (Kirchhoff et al., 2010). Geschlossene Fragen wurden so neutral wie möglich gestaltet, d. h. sie boten zwischen drei und sieben Antwortmöglichkeiten mit einem neutralen Mittelwert. Halboffene Fragen sollen die Vorzüge geschlossener und offener Frageformate vereinen und deren jeweilige Einschränkungen ausgleichen (Hollenberg, 2019). In dieser Variante wird das geschlossene Frageformat durch eine Restkategorie (z. B. „Sonstiges“) in Verbindung mit einem freien Kommentarfeld ergänzt (Hollenberg, 2019). Offene Fragen bieten einen freien Raum an Antwortmöglichkeiten zu den Befragten und werden individuell

bewertet. Der Fragebogen und die Art der Umfrage gewährleisten die Objektivität und Reliabilität der Untersuchung. Die Umfrage wurde mit Lamapoll (app.lamapoll.de) eines Berliner Softwareanbieters für Umfragen und Fragebögen im Online-Bereich erstellt. Die Methode der Online-Befragung wurde gewählt, weil auf diese Weise die befragten Personen schnell und einfach erreicht werden können. Außerdem kann mit dieser Methode die Anonymität der Teilnehmer gewährleistet werden. Die Teilnahme an der Umfrage ist sehr einfach und leicht durch das Versenden eines Links zur Umfrage, der von allen gängigen Endgeräten abrufbar ist. Eine Überprüfung der vorgesehenen Funktionalität des Fragebogens wurde vor der Durchführung der Befragung durch eine neutrale Person durchgeführt. Damit konnte insbesondere bei Online-Befragungen sichergestellt werden, dass keine inhaltlichen oder technischen Probleme auftreten.

6.1.2 Inhalt des Fragebogens

Im Hinblick auf die Forschungsfrage sollte die Auswahl und der Umfang des Fragenkatalogs einen möglichst gerechten Vergleich widerspiegeln. Darüber hinaus sollen die Antworten der Teilnehmer die aufgestellten Hypothesen möglichst genau entkräften oder belegen. Die konkrete Auswahl der Fragen orientiert sich an den Eigenschaften der eingesetzten Technik und der herangezogenen Fachliteratur aus den Bereichen der Medien- und Veranstaltungstechnik.

Nachfolgend werden die ausgewählten Bewertungskriterien und die gewählten Bewertungsmethoden für die ausgewählten Fragen erläutert. Eine vollständige Übersicht über alle Fragen ist im Anhang dargestellt.

Frage 1

„Bitte wählen Sie Ihr Geschlecht aus“

Durch die Unterteilung nach Geschlecht soll die Auswertung der gesammelten Daten übersichtlicher und spezifischer werden. Die Frage dient auch dazu, herauszufinden, ob es einen Zusammenhang zwischen dem Auftreten bestimmter VR-Krankheiten und dem Geschlecht gibt. Hier wird bei Geschlecht, zwischen „Frau“, „Mann“, „Divers“ und „Keine Antwort“ unterschieden.

Frage 2

„Zu welcher Altersgruppe gehören Sie?“

Durch die Unterteilung nach Altersgruppe soll die Auswertung der gesammelten Daten übersichtlicher werden. Die Frage dient auch dazu, herauszufinden, ob es einen Zusammenhang zwischen dem Auftreten bestimmter VR-Krankheiten und Altersgruppe Kategorie gibt. Hier wird die Altersgruppe in vier Kategorien aufgeteilt.

Frage 3

„Welches Equipment haben Sie verwendet, um das 360°-Video anzuschauen? (Mehrfachantworten möglich)“

Hier wird eine Frage zu dem verwendeten Equipment während der Befragung formuliert, die Frage dient dazu, die Relevanz der Studie zu überprüfen, denn für ein authentisches VR-Erlebnis wird eine VR-Brille benötigt. Da aktuell nicht viele Menschen die Möglichkeit haben, VR-Brillen zu nutzen, die meisten jedoch einfach, mit einem Computer, Smartphone oder Tablet, ein 360°-Video anschauen bzw. rotieren können, um die Perspektive zu wechseln, wurde dies berücksichtigt. Für die Bewertung des 3D-Sounds ist die Verwendung geeigneter Kopfhörer ein wichtiger Faktor, daher wurde ein Kommentarfeld angefügt, in dem die*der Teilnehmer*in die Art des verwendeten Kopfhörers beschreiben kann. Dies dient dazu, die Auswertung der Ergebnisse relevant und sinnvoll zu gestalten.

Frage 4

„Sind Sie schon einmal mit Virtual Reality oder 360°-Videos in Kontakt gekommen?“

Hier wird die Erfahrung der Teilnehmer*innen mit Virtual Reality oder 360°-Videos abgefragt, um später eine Vorstellung davon zu bekommen, wie hoch der Anteil der Personen ist, die bereits mit dieser Technologie in Berührung gekommen sind. Diese Frage dient auch dazu, herauszufinden, wie weitverbreitet diese Technologie für den Standardanwender ist.

Frage 5

„Wie fanden Sie die Bildqualität der 360°-Videoaufnahme im Vergleich zur 2D-Videoaufnahme?“

Diese Frage dient dazu, einen direkten Vergleich der Bildqualität und Auflösung der 360°-Videoaufnahme mit der 2D-Videoaufnahme zu machen, dafür wird eine fünfstufige Bewertung Skala von „viel schlechter“ bis „viel besser“ vorgegeben.

Frage 6

„Wie fanden Sie die 3D-Soundqualität der 360°-Videoaufnahme im Vergleich zu 2D-Videoaufnahme?“

Diese Frage dient dazu, einen direkten Vergleich der Soundqualität bzw. 3D-Ambisonic -Ton der 360°-Videoaufnahme mit der Stereo-Ton der 2D-Videoaufnahme zu machen. Auch hier wurde eine fünfstufige Bewertung Skala von „viel schlechter“ bis „viel besser“ vorgegeben.

Frage 7

„Wie sehr haben Sie sich beim Betrachten des 360°-Videos im Vergleich zu 2D-Aufnahme visuell umhüllt gefühlt?“

Diese Frage zielt darauf ab, das Gefühl der visuellen Umhüllung der 360°-Aufnahme direkt mit der 2D-Aufnahme zu vergleichen. Dies geschieht nur, wenn die*der Teilnehmer*in eine VR-Brille verwendet hat. Für diese Frage wird eine fünfstufige Bewertung Skala von „Gar nicht“ bis „Komplett umhüllt“ vorgegeben.

Frage 8

„Wie sehr haben Sie sich beim Anhören des Sounds des 360°-Videos im Vergleich zur 2D-Aufnahme auditiv umhüllt gefühlt?“

Diese Frage dient dazu, das Gefühl der auditiven Umhüllung der 3D-Ton direkt mit dem Stereo-Ton zu vergleichen. Dies geschieht nur, wenn die*der Teilnehmer*in geeignete

Kopfhörer verwendet hat. Für diese Frage wird eine fünfstufige Bewertung Skala von „Gar nicht“ bis „Komplett umhüllt“ vorgegeben.

Frage 9

„Wie stark war das Gefühl vor Ort zu sein bei der 360°-Videoaufnahme?“

Diese Frage bezieht sich auf das Gefühl, vor Ort zu sein oder Präsenz, d. h. ob die*der Teilnehmer*in das Gefühl hatte, in der virtuellen Welt dabei zu sein bzw. mit dieser interagieren zu können. Hier wird die*der Proband*in nach der Intensität dieses Gefühls gefragt. Die Bewertung erfolgt auf einer fünfstufigen Skala von „gar nicht“ bis „sehr stark“.

Frage 10

„Wie stark war das Gefühl vor Ort zu sein bei der 2D-Videoaufnahme?“

Bei dieser Frage wird die*der Teilnehmer*in gefragt, wie stark das Gefühl vor Ort zu sein ist, während des Betrachtens der 2D-Aufnahme. Die Bewertung erfolgt auf einer fünfstufigen Skala von „gar nicht“ bis „sehr stark“.

Frage 11

„Zurzeit kann man ein gutes VR-Headset, wie z. B. das Oculus Quest 2, für 349 € bestellen. Würden Sie sich für diesen Preis ein VR-Headset kaufen?“

Diese Frage untersucht den wirtschaftlichen Aspekt. Hierbei geht es um die Kaufbereitschaft der Teilnehmer*innen oder Verbraucher. Durch diese Frage wird bewertet, ob der Verbraucher den aktuellen Preis der VR-Brille ausgeben möchte, um sich das Gerät zu leisten. Zur Beurteilung der Kaufbereitschaft wurde eine 3-stufige Skala erstellt, (Ja, Nein, Vielleicht).

Frage 12

„Wie viel mehr (oder auch weniger) wären Sie bereit, für eine Online-Veranstaltung in gezeigten 3D-VR-Variante zu bezahlen (z. B. einen virtuellen Theater- oder Konzertbesuch)?“

Angesichts des finanziellen Aufwands zur Umsetzung einer Online-VR-Veranstaltung geht es hier um die Frage, ob die*der Teilnehmer*in bereit ist, mehr oder weniger Geld zu zahlen, um Online-Veranstaltungen im 360°-Format anzuschauen. Hier wird untersucht, ob es überhaupt sinnvoll ist, diese Technologie in der Veranstaltungsbranche einzusetzen. Um das Interesse der Teilnehmer*innen an diesem Aufnahmeformat zu bewerten, wird hier eine siebenstufige Skala erstellt, von (-10 €) bis (10 €).

Frage 13

„Haben Sie sich während des Betrachtens des VR-Videos unwohl gefühlt?
Wenn ja, wählen Sie bitte die passende Beschreibung (Mehrfachantworten möglich)“

Diese Frage untersucht den Komfort-Aspekt. Hier wird gefragt, ob sich die*der Teilnehmer*in beim Betrachten von 360°-Videos mit VR-Brillen unwohl gefühlt hat. Dafür gibt es verschiedene Gründe, denn laut Studien gibt es unterschiedliche Einflussfaktoren, die beim Betrachten auftreten können: Flimmern, die Bildrate, Tracking-Fehler, die Qualität der Systemkalibrierung und die Verwendung von stereoskopischen Bildschirmen können dazu führen, dass unerwünschte Nebeneffekte während des Betrachtens entstehen, wie z.B. Übelkeit, Schwindel, Hörstörung, Kopfschmerzen, usw. (Dörner et al., 2019) Da auch das einfache Tragen einer VR-Brille für den Betrachter störend sein kann, sind diese Aspekte für die Studie von großer Bedeutung. Um den Komfort und der ergonomische Aspekt zu bewerten, wurden sechs Antwortmöglichkeiten mit den am häufigsten auftretenden Effekten wie „Übelkeit“, „Schwindel“, „Unwohlsein“, „Sehstörung“, „Hörstörung“ formuliert. Außerdem wurde ein Kommentarfeld eingefügt, um weitere Nebenwirkungen zu nennen, die neue Informationen für die Auswertung der Ergebnisse liefern können.

Frage 14

„Bitte kommentieren Sie hier, was Ihnen im Zusammenhang mit Ihren abgegebenen Bewertungen evtl. aufgefallen ist/was Sie uns näher erläutern bzw. mitteilen wollen“

Diese offene Frage dient dazu, die*dem Teilnehmer*in einen Freiraum zu geben, um, wenn möglich, andere Aspekte der Forschung zu nennen, die vorher bei der Entwicklung des Fragebogens nicht erwähnt wurden. Für die Auswertung wird hier ein offenes Kommentarfeld angefügt.

6.2 Praktische Vorgehensweise

Für die praktische Umsetzung der Methodik wurde ein vierstufiges Vorgehen praktiziert: Probenaufnahme Immergut Festival (Immergutrocken, 2021), Kontaktaufnahme mit Theatern in Berlin und Planung der Aufnahme im Bluemax Theater Berlin, Durchführung der Echtzeit Parallelaufnahme mit der Blue Group im Bluemax Theater, Postproduktion und Online-Veröffentlichung.

Die erforderliche Ausstattung für die VR-Aufnahmen wurde von der Berliner Hochschule für Technik (BHT) bereitgestellt. Nachdem die Geräte ausgeliehen und voreingestellt waren, wurde die erste Testaufnahme durchgeführt. Bei der ersten Testaufnahme handelt es sich um eine klassische Open-Air-Konzertveranstaltung mit ca. 200 Zuschauern*innen. Ziel dieses Versuchs ist es, die vorhandene Technik zu testen, die Bereitschaft der Künstler zur Nutzung von neuen Technologien zu prüfen und das Publikum mit dieser neuen Art von Aufnahmen zu konfrontieren. Und anschließend die Qualität und das Erlebnis der beiden Aufnahmen zu vergleichen, um mögliche Fehler im Produktionsprozess zu korrigieren, damit die Umsetzung der bevorstehenden Theateraufnahme optimiert werden kann.

Nach der Durchführung der ersten Probeaufnahme, wurden zunächst verschiedene Berliner Theater kontaktiert, um zum einen die bisherigen Erfahrungen mit VR-Aufnahmen im professionellen Umfeld umzusetzen und erfragen, ob ein Interesse gibt VR- bzw. 360°-Videoaufnahme in den Berliner Theatern durchzusetzen.

Die Durchführung der Echtzeit VR Aufnahme erfolgte mithilfe eines Assistenten, Herr Mohamed Amine Mechria, der für die 2D-Kamera zuständig war. Die Kameras wurden

in zwei verschiedenen Reihen platziert. Mit den Erkenntnissen des ersten Probeversuchs wurden zusätzliche Verbesserungen bei der Gestaltung und Einstellung des Materials vorgenommen (die genaue Vorgehensweise wird in Kapitel 6.2.2 erläutert). Schließlich erfolgt in der Postproduktion die Bearbeitung der Videos mit Adobe Premier Pro. Für 3D-Audio wurde die kostenlose DAW Reaper verwendet. Die fertige Produktion wurde anschließend in YouTube hochgeladen. Anschließend erfolgte eine Bewertung durch 15 Personen in Form einer Umfrage.

Material Einstellung

Sony PXW-X70:

Die Standardvoreinstellungen der Kamera unterstützen die maximale 1080p/60p -Auflösung. Für 4K Aufnahmen musste der Firmware auf die Version 3.0 oder aktualisiert werden (Sony, 2022). Nach der Aktualisierung der Kamera direkt via WLAN-Verbindung über das wurde die Aufzeichnung in 4K (2160p/30p) -Auflösung angezeigt (siehe Abb. 26) und dann gewählt. Ziel ist es, 4K-Auflösung für beide Testkameras bereitzustellen, um einen fairen Vergleich der Bildqualität zu ermöglichen.

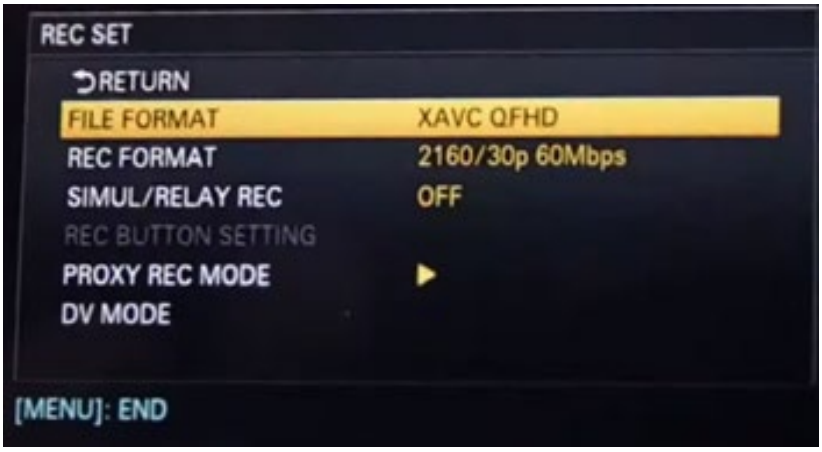


Abb. 26: Aufnahmeeinstellungen Sony PXW-X70 Kamera

Weitere Einstellungen wie Weißabgleich, Kontrast und ISO wurden kurz vor der Aufnahme angepasst.

Vuze 3D 360° 4K

Die Vuze-Kamera hat kein Display und kann entweder über die beiden Tasten mit farbigen LED-Anzeigen (Power, Aufnahme), die sich auf der Oberseite der Kamera befinden (vgl. Abb. 28), oder über die Vuze Camera App bedient werden. Möchte der Nutzer die Videoeinstellungen anpassen, geschieht dies ausschließlich über die App. Jedes der einzelnen Sony-Objektive der Vuze kann UHD 4K-Videos mit 30p aufzeichnen (Humaneyes Technologies, 2017). Die Kamera speichert die Daten auf einer microSD-Karte in einem externen Speicherslot an der Seite der Kamera, der sich zwischen den beiden Objektiven befindet. Das Gerät wiegt etwa. 450 g und ist nach IP64 vor Spritzwasser und Staub geschützt (Humaneyes Technologies, 2017). Außerdem verfügt die Vuze-Kamera über vier integrierte Mikrofone für Stereo-Audioaufnahmen (vgl. Abb. 27). Das vollständige Aufladen der Kamera dauert 3 Stunden bei einer Betriebsdauer von 2 Stunden. Die Funktionen Fokus und Zoom sind bei dieser Kamera nicht verfügbar (Humaneyes Technologies, 2017).

Abb. 27: Eigenschaften Vuze 3D-360-Kamera (Humaneyes Technologies, 2017)

Abb. 28: Statusanzeigen Vuze 3D-360-Kamera (Humaneyes Technologies, 2017)

Das Gerät musste zuerst aktualisiert werden, damit es von dem verwendeten Smartphone (Samsung S8) erkannt wurde. Nachdem die Kamera aktualisiert war, konnte sie über WLAN an dem Smartphone verbunden werden. Sobald die Verbindung zum Smartphone hergestellt ist, wird eine Live-Vorschau der Szene in allen vier Richtungen auf dem Gerät angezeigt. Über die App besteht die Möglichkeit, verschiedene Parameter einzustellen, man kann z. B. die Helligkeit der einzelnen Objektive anpassen, die Kamera

senkrecht kalibrieren, die Videoaufnahmequalität verbessern und die SD-Karte formatieren.

Über die Vuze-Camera-App wurden verschiedene Einstellungen vorgenommen, um die beste Aufnahmequalität herauszuholen. Hier wurden die Bitrate, 120 Mbps, Antiflicker bei 60 Hz und der Videomodus 3D-30fps eingestellt (vgl. Abb. 29).

Abb. 29: Aufnahmeeinstellungen Vuze 3D-360-Kamera (Bildschirmfoto, Vuze Camera app)

Nach der Einstellung aller Parameter kann der Nutzer entweder den Foto- oder den Videomodus auswählen und die Aufnahme starten, indem er die Aufnahmetaste über die App oder direkt über die Aufnahmetaste der Kamera drückt. Zu Beginn der Aufnahme ertönt ein einzelner Piepton, am Ende der Aufnahme ertönen zwei Pieptöne. Die Bild-/Videodateien sind im Bibliotheksbereich der App oder durch Verbinden der Kamera mit einem Micro-USB-Kabel an den Computer zugänglich.

ZOOM F6 Rekorder

Das F6 bietet die Möglichkeit, im 16/24-Bit-WAV- und im 32-Bit-Float-WAV-Format aufzunehmen. Bei einer Aufnahme im 32-Bit-Float-WAV-Format bleibt die Auflösung der Aufnahme auch nach einer nachträglichen massiven Änderung der Pegel erhalten (ZOOM CORPORATION, 2022). Das Gerät bietet außerdem die gleichzeitige Aufnahme von 6 Kanälen und 14 Spuren (ZOOM CORPORATION, 2022). Es unterstützt drei Batterietypen (USB-Batterie, eine L-Batterie und Batterien vom Typ AA) (ZOOM CORPORATION, 2022). Der Recorder verfügt über ein Display und kann entweder direkt über die Bedientasten bedient werden oder durch die Installation eines ZOOM

Drahtlosadapters (z. B. BTA-1) ist eine Fernsteuerung und damit die Nutzung der iOS-App F6 Control möglich (ZOOM CORPORATION, 2022). In Kombination mit VR-Mikrofonen ermöglicht der Ambisonic-Modus eine räumliche 360°-Audioaufnahme (ZOOM CORPORATION, 2022). Neben Link-Funktionen für Pegelregler- und Einstellungen wird auch die Dekodierung von Ambisonic-Format A nach Format B unterstützt (ZOOM CORPORATION, 2022).

Vor der ersten Probeaufnahme beim Immergut Festival wurden folgende Einstellungen zur Freischaltung des Ambisonic-Modus vorgenommen:

Zuerst wurde der duale Aufnahme-Modus ausgewählt. In diesen Modus werden gleichzeitig herkömmliche 16/24 Bit WAV-Dateien und 32 Bit Float WAV-Dateien aufgenommen. Die Samplerate 48kHz wurde als Standard für diese Aufnahme ebenfalls gewählt. Das Dateiformat Poly wurde ausgewählt, um eine einzelne Mehrspur-Datei mit mehreren Audiospuren zu erstellen.

Zur Freischaltung des Ambisonic-Modus wurden folgenden Schritte vorgenommen:

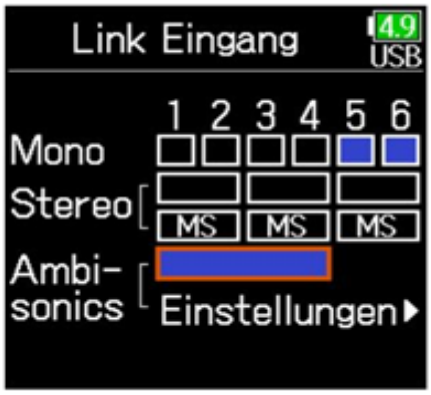


Abb. 30: Einstellung des Ambisonic-Formats in ZOOM F6 Rekorder (ZOOM CORPORATION, 2022)

Über die Menütaste des Rekorders wurde die Ambisonics-Funktion aktiviert (vgl. Abb. 30). Die Link-Einstellung dient dazu, die Kanäle 1 bis 4 mit einer einzigen Ebenentaste zu verknüpfen, dadurch können die 4 Kanäle mit einer einzigen Taste bedient werden. Dies sorgt dafür, eine gleiche Laustärke in die verwendeten 4 kanälen zu gewährleisten, sodass der erste Kanal bei der Codierung der vier Kanäle von keinem anderen Kanal im Pegel überschritten wird.

6.2.1 Durchführung einer Probe/Testaufnahme am Immergut Festival

Für die Probeaufnahme auf dem Festival wurde folgende Technik eingesetzt:

360°-Kamera	Vuze 3D 360 4K VR Camera
2D-Kamera	Sony PXW-X70
3D AUDIO Mikrofon	Sennheiser AMBEO VR Mic
Rekorder	Zoom F6
1x Handy für die Fernsteuerung	Samsung S8, iPhone 6

Die Technik wurde dann schrittweise eingestellt, um die bestmögliche Kapazität der Geräte zu gewährleisten.

Im Rahmen dieses Projekts wurden sechs verschiedene Bands kontaktiert, von denen sich drei zurückmeldeten und zusagten. Eine Band sagte das Projekt aufgrund der kurzfristigen Kontaktaufnahme ab und zwei Bands meldeten sich nicht zurück. Als Ergebnis der Kontaktaufnahme kann man feststellen, dass es bereits ein Interesse an der Nutzung von VR-Aufnahmen auf Seite der Musiker besteht. Aufgrund des Regens und der schlechten Wetterbedingungen konnte nur eine einstündige Konzertaufnahme angefertigt werden.

Die erste Probeaufnahme fand am 27. August 2021 in Neustrelitz mit der Berliner Indie-Pop-Band THALA statt (THALA, 2022). Für diesen ersten Probeversuch wurde das Material im Vorfeld gemäß dem Abschnitt Material-Einstellung eingestellt. Als Aufnahmeposition für die 360°-Kamera und die 2D-Kameras wurde die vordere Mitte des Publikumsbereichs gewählt, in einer Entfernung zur Bühne von etwa 5 m. Die Kameras wurden parallel zueinander mit einem Abstand von ca. 1 m, auf Augenhöhe von ca. 1,70 m, platziert, um den Blick einer „stillerstehender Mensch“ im Publikum zu simulieren (Windsheid und Gold, 2022).

Die 360°-Kamera wurde auf einem Stativ (BENRO A2573F) montiert und die 2D-Kamera auf einem Stehtisch in der gleichen Höhe aufgestellt. Aufgrund fehlender passender Mikrofonhalterung für das Stativ wurde das 3D-Ambeo-Mikrofon an der Sony-Kamera in Aufwärts (eng. Endfire) -Position aufgesetzt, welche bereits über eine eingebaute Mikrofonhalterung verfügt. Die Aufbauzeit betrug ca. 15 Minuten.

Bei teilweisem Regenwetter wurde ein Regenschirm verwendet, um die Geräte zu schützen. Der Regen zu Beginn des Konzertes war glücklicherweise harmlos und dauerte nur 10 Minuten. Zu Beginn des Konzerts. Die Videoaufnahmen wurden synchron durchgeführt in Zusammenarbeit mit zwei weiteren Studenten der BHT, Herr Sebastian Finsterer und Herr Marius Mirschberger, um den reibungslosen Ablauf des Versuchs zu gewährleisten.

Ziel dieses ersten Versuchs war es erstens, das Equipment zu testen, um sicherzustellen, dass alle Geräte ordnungsgemäß funktionieren, um später eventuell Anpassungen an den Einstellungen des Materials sowie bei der Gestaltung vorzunehmen. Zweitens sollte für den Test das Blickfeld einer stillstehenden Person im Publikum simuliert werden, um sicherzustellen, dass die virtuellen Zuschauer*innen das Gefühl haben, „Teil des Geschehens“ zu sein und das Konzert so zu verfolgen, als wären sie selbst Teil der Veranstaltung (Windsheid und Gold, 2022).

Beurteilung und Optimierung

Nach der Durchführung des ersten Probeversuchs wurden die gespeicherten Rohdaten in den Rechner importiert und mithilfe eines kostenfreien VR-Videoplayers DeoVR (STEAM, 2022) in der SteamVR Plattform mit der VR-Brille visualisiert. Die Ergebnisse der Visualisierung waren einigermaßen zufriedenstellend, da die Simulation-aufnahme gut zu erkennen war. Die Bildqualität war ebenfalls akzeptabel, da die Aufnahme bei Tageslicht durchgeführt wurde. Bei der 2D-Aufnahme waren die Visualisierungsergebnisse nicht überzeugend, da die Kamera als Zwischenlösung auf einem Tisch platziert wurde, was keine gute Entscheidung war, da der Tisch in bestimmten Momenten der Aufnahme wackelig war, was die Stabilität der

Videoaufnahme und folglich auch der 3D-Audioaufzeichnung mit einigen Kratzgeräuschen beeinträchtigte hat. Außerdem fiel der Rekorder zweimal wegen eines wackeligen USB-C-Kabels aus, das zur Stromversorgung des Rekorders über die Powerbank eingesetzt wurde.

Einer der auffälligsten Punkte bei der Optimierung der Gestaltung für die bevorstehende Aufnahme war die Bereitstellung eines zweiten Stativs für die zweite Kamera und eines geeigneten Mikrofonhalters. Ebenso wichtig war die Anwendung der Fernsteuerung über die iOS-App ZOOM, die den Betriebsstatus des Rekorders in Echtzeit anzeigt. Daher war es notwendig, ein zweites iOS-Handy zu verwenden, da die App nicht mit Android-Geräten kompatibel ist. Darüber hinaus wurde beschlossen, zusätzliche Batterien zu verwenden, um den ordnungsgemäßen Betrieb des Rekorders zu gewährleisten.

Daraus resultierend werden für die kommende Aufnahme die folgenden Änderungen vorgenommen: Erstens wird ein zweites Stativ zur Gewährleistung einer stabilen Aufnahme eingesetzt, die Stromversorgung der Rekorder wird durch zusätzliche Batterien sowie die Powerbank sichergestellt, was einen Ausfall der Rekorder verhindern soll. Zur Optimierung der Tonaufnahme wird das Mikrofon unter der VR-Kamera in Aufwärtsrichtung (eng. Upright) Position mit einer geeigneten Mikrofonhalterung angebracht, um die Reflexion des Bodens zu reduzieren und einen authentischen Eindruck der Umgebung zu gewährleisten.

6.2.2 Durführung der Echtzeit Aufnahme im Bluemax Theater

Während der Pandemiezeit wurde die gesamte Veranstaltungs-, Theater- und Kulturbranche hart getroffen. Zahlreiche Shows, Konzerte und Aufführungen wurden abgesagt oder verschoben. Der Veranstaltungssektor blieb jedoch nicht untätig, einige Proben fanden weiterhin in Theatern statt, wobei strengere Hygieneregeln eingehalten werden mussten. Andere Veranstaltungen wiederum fanden im Internet in Form von Online-Streaming-Veranstaltungen statt, da Online-Kanäle zu dieser Zeit die einzige Möglichkeit waren, ein breites Publikum zu erreichen.

Mit dem Sinken der Covid-Inzidenzwerte konnten wieder Veranstaltungen stattfinden, unter Einhaltung besonderer Hygienekonzepte stattfinden, um das Ansteckungsrisiko zu reduzieren. In dieser Phase wurde die Kontaktaufnahme mit den Berliner Theatern initiiert. Insgesamt wurden 18 Theaterhäuser per E-Mail oder Telefon kontaktiert. 12 davon antworteten und lehnten die Teilnahme an dem Masterprojekt ab. 5 Anfragen blieben ohne Rückmeldung. Ein Theater sagte seine Teilnahme an dem Projekt am 28. November 2021 zu. Das Bluemax Theater Berlin, das seit 15 Jahren die Show der Blue Man Group aufführt.

Nach der erfolgreichen Kontaktaufnahme lud mich der Theater Leiter Herr Stefan Meinhold ins Bluemax Theater ein. So konnte ich mir vorab die Show anschauen und die Show-Crew, den künstlerischen Leiter und den Blue-Man-Darsteller Herr Nadim Helow kennenlernen. Des Weiteren konnte ich bei dem Besuch in Zusammenarbeit mit dem technischen Leiter und Herr Helow die Kamera-Positionen für die Aufnahme festlegen. Nach dem Anschauen der Theatershow wurde dem künstlerischen Leiter die Positionierung der Kameras übermittelt: Beide Kameras sollten in der Mitte des Zuschauerraums platziert werden. Es wurde ferner verabredet, dass für den Aufbau des Aufnahme-Materials die Pause genutzt werden sollte, um die letzten 30 Minuten aufzunehmen, da die Lichteffekte zu diesem Zeitpunkt am intensivsten sind und dies für eine gelungene 360°-Aufnahme entscheidend ist. Herr Helow leitete meinen Vorschlag an den Theaterleiter und den Technischen Leiter weiter. Eine Woche später wurde der 9. Dezember 2021 für die Durchführung der Echtzeit-Aufnahme ausgewählt.

Mit den Erkenntnissen des ersten Versuchs wurden die Ursachen für die verbesserungswürdigen Fehler in der Methodik ermittelt und Optimierungsschritte für die Durchführung der Echtzeitaufnahme vorgenommen. Der technische Versuchsaufbau wurde am 09.12.21 im Bluemax Theater 30 Minuten vor Beginn der Show umgesetzt. Eine Checkliste diente dazu, die notwendigen Vorbereitungsschritte wie das Laden der Akkus, die Reinigung des Kameraobjektivs und das Formatieren der SD-Karten zusammenzufassen.

Das verwendete Equipment für den Probeaufbau ist in der folgenden Tabelle aufgelistet:

360°-Kamera	Vuze 3D 360 4K VR Camera
2D-Kamera	Sony PXW-X70
3D AUDIO Mikrofon	Sennheiser AMBEO VR Mic
Rekorder	Zoom F6
2x Handy zur Fernsteuerung	Samsung S8, iPhone 6
Kopfhörer	Sony MDR-7506
2x Stativ	BENRO A2573F, Rollei C5i

Für diesen technischen Probeaufbau wurden zwei verschiedene Aufnahme-Position im Saal gewählt, welche in Abbildung 31 markiert sind.

Abb. 31: Kamerapotionen im Zuschauerbereich während der Echtzeit-Aufnahme im Bluemax Theater Berlin, (1) Vuze 3D 360 4K VR Kamera und Sennheiser AMBEO VR Mikrofon, (2) Sony PXW-X70 Kamera (Sghaier, 2022)

Während des Probenaufbaus wurden die optimalen Standorte für die Kameras im Zuschauerraum festgelegt, die Aufnahmepositionen wurden unter Berücksichtigung der Anweisungen des Technischen Leiters gewählt. So wurde darauf geachtet, die Kameras

nicht ganz vorne auf der Bühne zu platzieren, da während der Aufführung Wasser auf das Publikum gespritzt wird und die Technik dadurch hätte beschädigt werden können. Die Kameras durften für das Publikum kaum sichtbar sein.

Vor dem Beginn der Aufführung wurde das Publikum darüber informiert, dass im zweiten Teil der Show eine Aufnahme für ein Masterprojekt stattfinden sollte, und alle befragten Personen waren damit einverstanden. Anschließend konnten die Kameras an den angegebenen Stellen platziert werden. Die 360°-Kamera wurde in der Mitte des Publikums in der 6. Reihe, Sitznummer 5, auf Augenhöhe ca. 1,60 m platziert, um die Wahrnehmung einer "stillsitzender Mensch" vor Ort zu simulieren. Das Ambeo-Mikrofon wurde mit Hilfe der Mikrofonhalterung im Stativ direkt unter der Vuze-Kamera in aufrechter (Upright) Position befestigt, der Rekorder in das Stativ eingehängt. Zum Abhören des Live-Sounds wurde ein Kopfhörer direkt an den Recorder angeschlossen (siehe Abb.32).

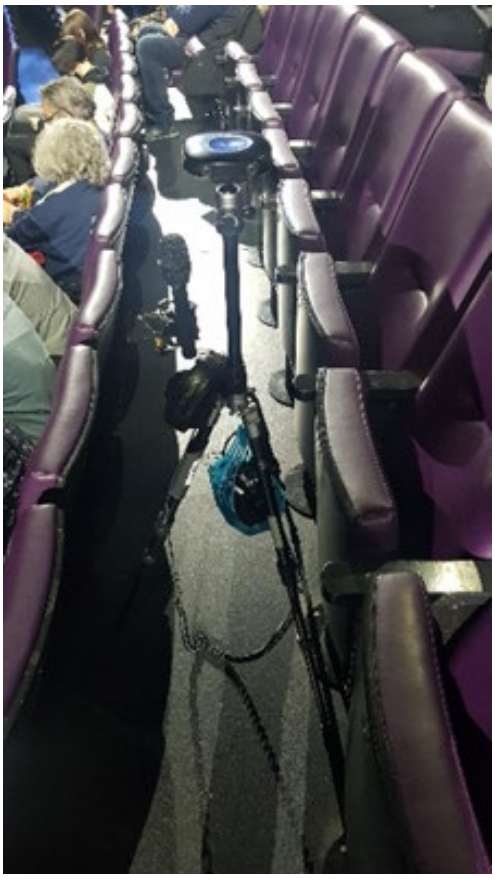


Abb. 32: Vuze 3D 360 4K VR Kamera, Sennheiser AMBEO VR Mikrofon und Zoom F6 Rekorder im Bluemax Theater Berlin (Sghaier, 2022)

Abb. 33: Sony PXW-X70 Kamera im Bluemax Theater Berlin (Sghaier, 2022)

Die zweite Kamera (siehe Abb.33) wurde weiter hinten, in der Mitte der 9. Reihe, in Augenhöhe angebracht. Das Ziel der Videoaufname war es, die gleiche Betrachtungsperspektive auf beiden Kameras zu erhalten. Das resultierende Bild der parallelen Aufnahme ist in Abbildung 34 zu sehen.

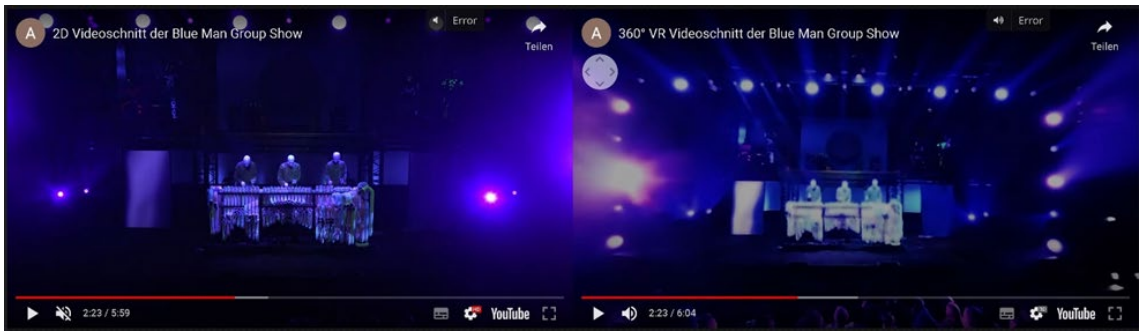


Abb. 34: Parallelaufnahme Übersicht der Blue Man Group Show 2021 im Bluemax Theater Berlin, links ist die 2D-Videoaufnahme und rechts die 360°-Videoaufnahme zu sehen (YouTube, 2022)

Beurteilung

Mit den Erkenntnissen der ersten Probeaufnahme konnte die Echtzeit-Aufnahme im Bluemax Theater reibungslos durchgeführt werden. Die Verwendung eines zweiten Handys zur Fernsteuerung des Rekorders (siehe Abb. 35) sparte viel Zeit und Stress, da man einen direkten Überblick über den Status der Aufnahme hat. Die Checkliste diente dazu, Akku-, Karten- und Speicherprobleme zu verhindern.

Abb. 35: Fernsteuerung des Zoom F6 Rekoder durch die iOS-App F6 Control (Bildschirmfoto, iPhone 6)

6.2.3 Postproduktion

Videopostproduktion:

Zuerst müssen die gespeicherten Videodateien importiert werden. Sobald die Kamera via USB- Kabel mit dem Computer angeschlossen ist, wird die Vuze-VR-Studio-Software automatisch gestartet. Über die Software kann die Option „Import top preview and edit“ ausgewählt werden. Alle aufgezeichneten Videodateien auf der Micro-SD-Karte der Kamera werden angezeigt und man kann beliebige Daten auswählen, um sie an einem bestimmten Speicherort auf dem jeweiligen Rechner zu importieren. Wenn das Video importiert ist, kann die Option „Vorschau und Bearbeiten“ engl. „Preview and edit“ auswählt werden, um zusätzliche Änderungen an der Datei vorzunehmen.

Vuze VR Studio

Vuze VR Studio ist eine VR-Videoproduktionssoftware, die mit der VUZE-Kamera geliefert wird. Das Programm kann Vuze-Material sofort erfassen, importieren und bearbeiten. Das VR-Inhalt nach der Bearbeitung mit der Software kann auf jeder Plattform angeschaute oder geteilt werden. Die Software ist mit Windows und MAC kompatibel (Vuze, 2020). Es besteht aus drei Bereichen: „Der Medienbrowser“, „Wiedergabebereich“ und „Der Werkzeuge Bereich“ engl. Tool Box enthält mehrere Optionen, um den VR-Inhalt zu bearbeiten.

Für dieses Projekt wurden einige Werkzeuge verwendet, in den Output-Einstellungen wurde die Option Output Type 3D-360-Video gewählt, Seitenverhältnis 1: 1. Die Blickposition wurde mit der Option "Ajust FOV and Center" festgelegt. (siehe Abb. 36).

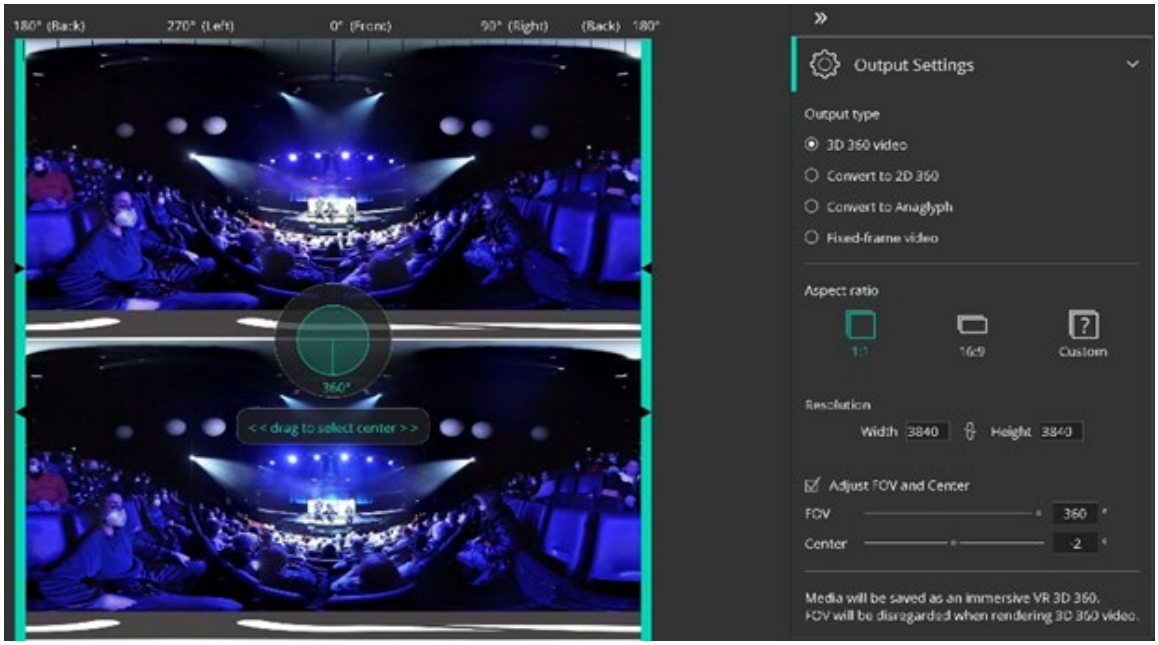


Abb. 36: Output-Settings Übersicht in Vuze VR Studio Software, 360°-Video der Blue Man Group Show 2021 im Bluemax Theate Berlin (Vuze VR Studio, 2020)

Mit der Option „Stitching“ (Deutsch: Zusammennähen) (Windscheid und Gold 2022). können die sich überlappenden Bereiche der verschiedenen Linsen zusammengefügt werden, so dass ein sphärisches 360°-Bild entsteht (Feyder, Rath-Wiggins, 2018). Dieser Prozess wurde automatisch über Vuze-Kamera durchgeführt. Und nochmal manuell im Detail über die Option „Refine Stitching“ angepasst (Vgl. Abb. 37, 38).

Abb. 37: 360°-Video der Blue Man Group Show 2021 im Bluemaxtheater Berlin vor der Verwendung Refine Stitching Option (Vuze VR Studio, 2020)

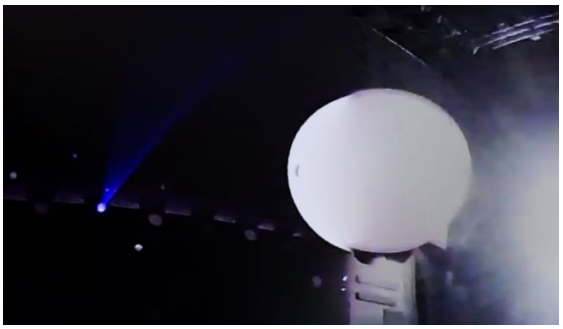


Abb. 38: Verwendung der Refine Stitching Option in Vuze VR Studio Software, 360°-Video der Blue Man Group Show 2021 im Bluemaxtheater Berlin (Vuze VR Studio, 2020)

Das Zoom-Tool hilft während des Stiches dabei, die Stichlinien genauer zu korrigieren. Das Rendering wurde als letztes durchgeführt. Dabei wurde der verarbeitete 360°-Inhalt

im H-264-Format enkodiert, für die Bitrate wurde die Option „Auto optimal“ gewählt, die Projektion wurde Equiranectangulare festgelegt, und Stereo-Audio (Vgl. Abb. 39). Da die Vuze-Kamera nur in Stereo aufzeichnen kann, wurde die 3D-Audiodatei später mit Adobe Premier Pro zum Videomaterial hinzugefügt. Dieser Prozess wird in späteren Abschnitten genauer erläutert.

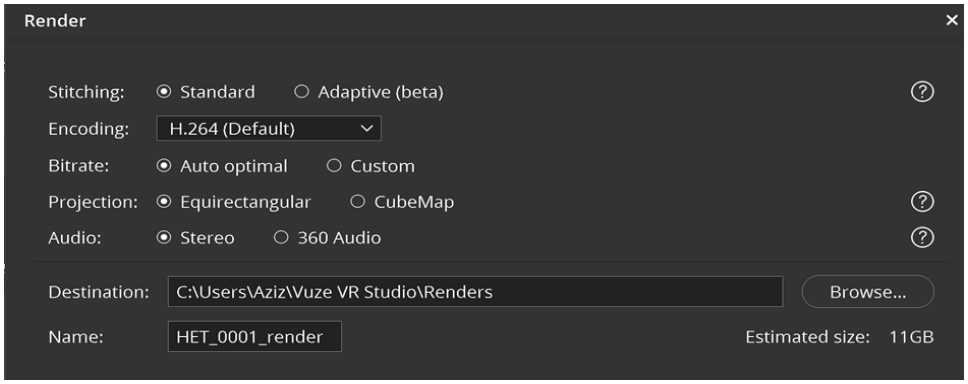


Abb.39: Eigenschaften der 360°-Videodatei (Vuze VR Studio, 2020)

Videomontage mit Adobe Premiere Pro

Nachdem das 360°-Video mit Vuze VR Studio bearbeitet und gerendert wurde, kann der VR-Inhalt in der Postproduktion mit einem geeigneteren Schnittprogramm bearbeitet werden. Zu diesem Zweck wurde eine von der Hochschule lizenzierte Schnittsoftware Abode Premiere Pro (Adobe, 2022) verwendet. Das Programm verfügt über viele VR-spezifische Optionen (siehe Abb.40). So können beispielsweise zusätzliche Grafiken oder Texte in das 360°-Format eingefügt werden. Die sphärischen Inhalte können direkt im Programm gedreht werden, außerdem können Überblendungen und Effekte ohne Verzerrung einfach eingebracht werden (Feyder und Rath-Wiggins, 2018). Die Gesamtdauer der Aufnahme beträgt etwa 25 Minuten, was für die Untersuchung zu lang ist. Daher wurden unnötige Passagen mit dem Programm entfernt und das Video um etwa 6 Minuten verkürzt. Derselbe Vorgang wurde für die 2D-Aufnahme wiederholt mit dem gleichen Programm.

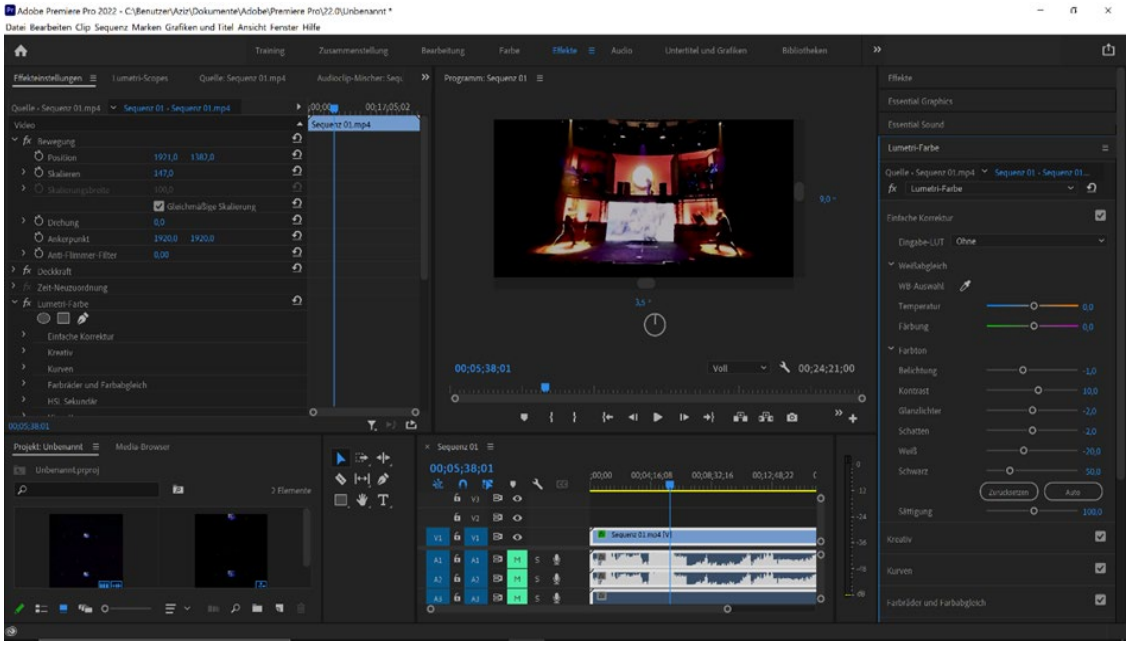


Abb. 40: 360°-Videomontage der Blue Man Group Show mit Adobe Premiere Pro (Adobe Premiere Pro, 2022)

Einige Farbkorrekturen wurden anhand der „Lumetri-Farbe“ in Adobe Premiere Pro vorgenommen, hier wurde hauptsächlich die Belichtung und Kontrast angepasst.

Tonbearbeitung

Für die Bearbeitung der 3D-Audioaufnahmen wurde die Audiosoftware Reaper (reaper.fm) verwendet, da sie Ambisonics höherer Ordnung unterstützt und die Möglichkeit bietet, bis zu 64 Spuren zu verwenden. Außerdem ist das Programm preisgünstig und bietet eine voll funktionsfähige kostenlose Testversion an (Zotter, Frank, 2019). Jede andere DAW, die VST und ausreichend viele Mehrspuren unterstützt (z. B. ProTools, Cubase) kann ebenfalls verwendet werden (Zotter, Frank, 2019).

Das Audiosignal wurde mit dem Ambeo Ambisonics-Encoder (Abb. 41) vom Ambisonics-A-Format in das B-Format konvertiert. Um sicherzustellen, dass das Audiosignal später von dem Streaming-Dienst YouTube erkannt wird, welcher als Plattform für die Veröffentlichung der Videos gewählt wurde. Als Output-Format wurde ambiX gewählt, das von den meisten Software und Plugins unterstützt wird.

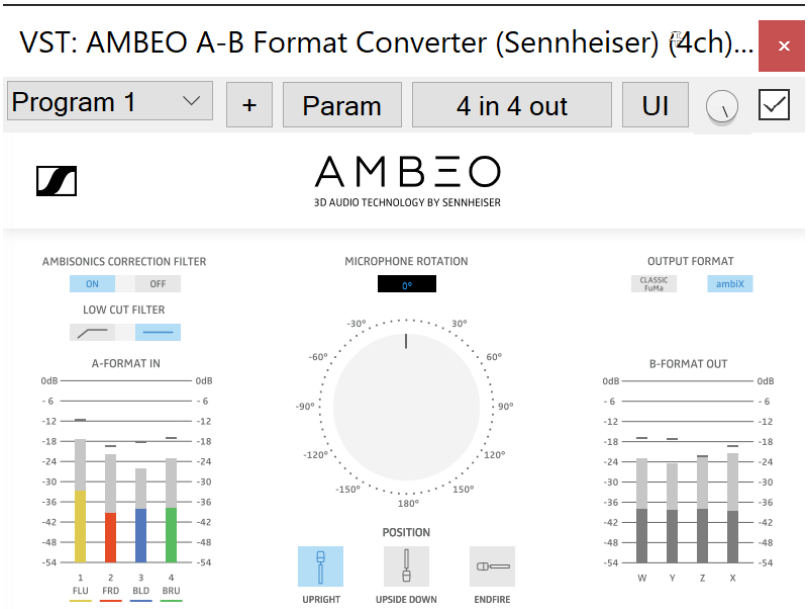


Abb.41: AMBEO A-B-Konverter in Reaper

Um die Hörrichtung zu visualisieren, wurde ein weiteres Plug-in aus der IEM Plug-in-Suite (IEM, 2021), der „EnergyVisualizer“ verwendet (siehe Abb.42).

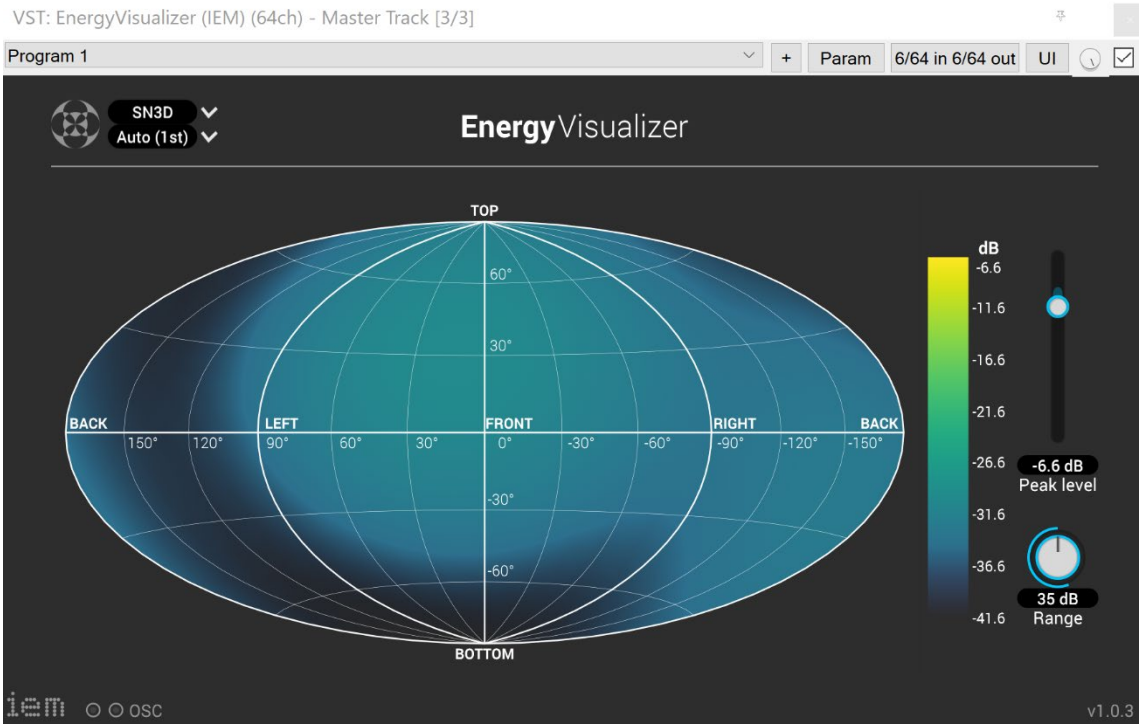


Abb. 42: EnergyVisualize in Reaper

Dieses Plug-in ist perfekt für die Messung des räumlichen Bildes geeignet. Der EnergyVisualizer visualisiert die Energieverteilung auf der Kugel des Ambisonic-

Eingangssignals. Die Energieniveaus sind entsprechend der Farbkarte farbkodiert Daher wurde dieses Plug-in in Kombination mit AMBEO A-B-Konverter verwendet, um den Fokus auf der Bühne (Front) auszurichten.

Zwei weitere Plugins der Softwaresuite von Facebook Spatial-Workstation wurden eingesetzt. Der FB360-Konverter (ambiX), (siehe Abb.43), ermöglicht eine Echtzeit binaurales Abhören des ambisonischen Signals, und das Audio 360 Meter (siehe Abb. 44) bietet einen Überblick über die Lautheit der gesamten räumlichen Mischung. Dieses Plug-in liefert nützliche Daten, die verhindern, dass der endgültig veröffentlichte Inhalt bei der Wiedergabe auf dem Zielgerät verzerrt wird. Das Audio 360 Meter wurde in die Master-Spur gelegt und die integrierte Lautheit eng. Loudness Units relative to Full Scale (LFUS) wurde gemessen, um den Standard-Youtube-LFUS-Wert von -14db (Tilgner, 2021) und dem wahren Spitzenwert engl. Decibels relative to full scale (dBFs) von -1db zu erzielen.

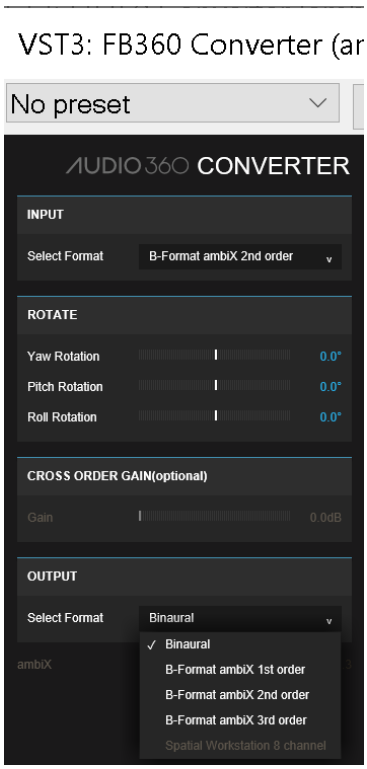


Abb. 43: Audio 360 Meter Plugin in Reaper

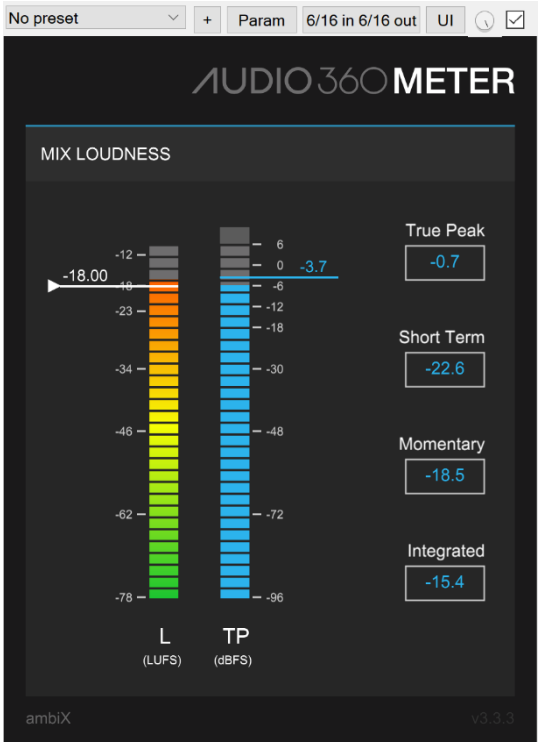


Abb. 44: Audio 360 Meter Plugin in Reaper

Export

Audioexport in Reaper:

Nach Abschluss der Audiobearbeitung wurde die Ambisonische Audiodatei mit folgende eingaben exportiert in Reaper (siehe Abb.45):

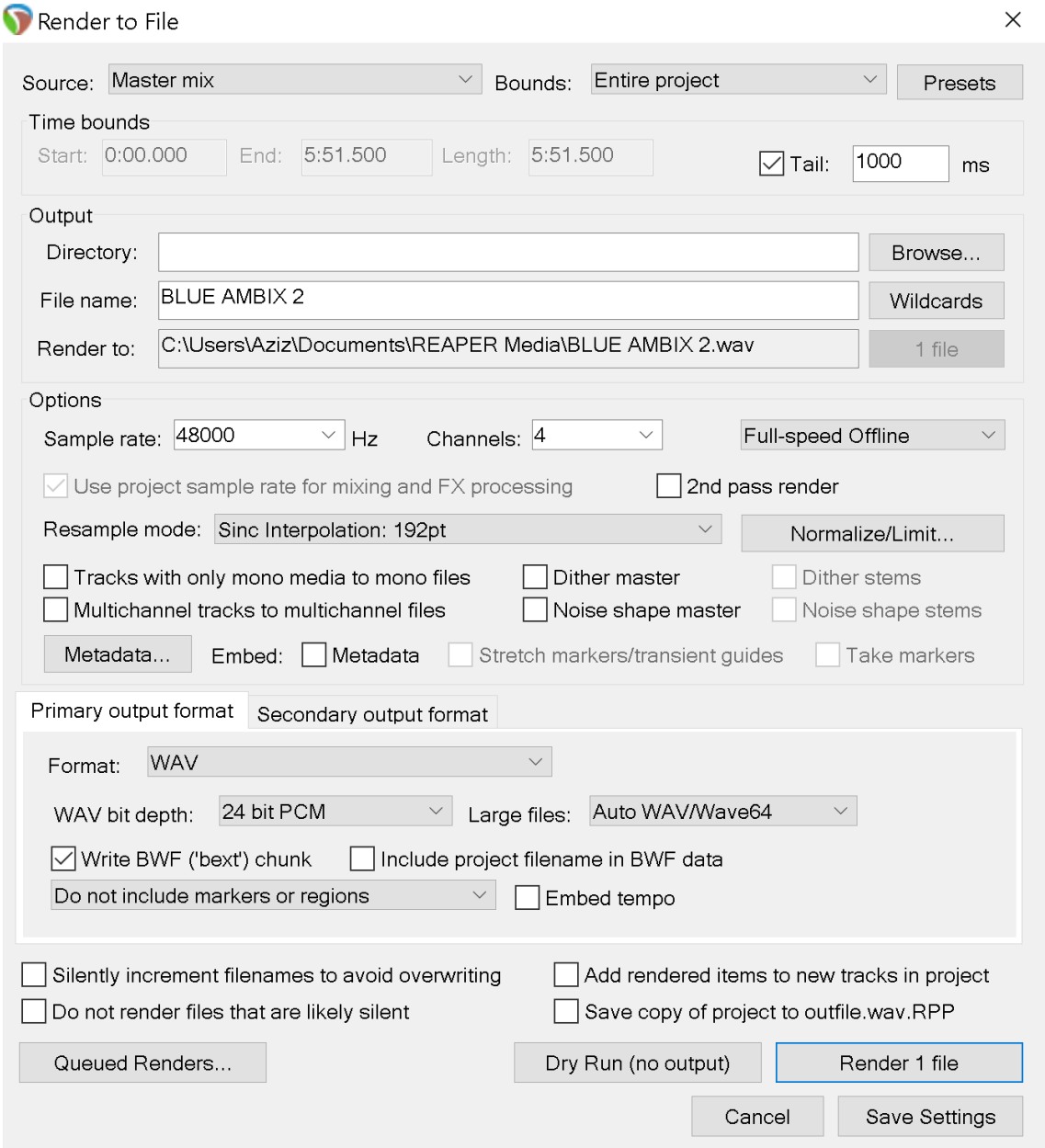


Abb.45: Export der 3D-Audiodatei in Reaper

Hier wurde die Option Channel von Stereo in eine 4-Kanal-AmbiX-Audiospur mit der Reihenfolge W, Y, Z, X geändert und die Abtastrate 48 kHz als Standard gewählt, da sie

von Streaming-Plattformen wie YouTube unterstützt wird (YouTube-Hilfe, 2022). Genau derselbe Rendering Prozess wurde für die 2D-Audiospur festgelegt, nur dass in diesem Fall die Audiospur im Stereoformat exportiert wurde.

Die Ambisonic- und Stereo-Audiodateien wurden in Adobe Premier Pro importiert, mit der Tonspur der eingebauten Lautsprecher beider Kameras mithilfe der Synchronisierungsoption in Adobe synchronisiert und schließlich ersetzt.

Videoexport in Adobe Premier Pro:

360°-Video export in Adobe Premier Pro:

Um 360°- Video in Adobe Premier Pro richtige zu exportieren sind folgende Schritte erforderlich:

1. Datei →Exportieren → Medien → Format: H.264 → Voreinstellung: Stereoscopic Over-Under Match Source Ambisonics → Bitrate 53-68 Mbps → Kontrollkästchen Video ist VR aktivieren.
2. Unter Audio → Audio Format AAC → Abtastrate 48.Hz → Kanale 4.0 → Bitrate Einstellungen 512 kbps → unter Ambisonics: Kontrollkästchen Audio ist Ambisonics aktivieren → Export. Ambisonic Kontrollkästchen erscheint nur, wenn das richtige Voreinstellung ausgewählt ist.

Für das Exportieren des 360°-Video wurde die maximale Renderqualität eingesetzt.

2D-Video export in Adobe Premier Pro:

Datei →Exportieren → Medien → Format: H.264 → Standard: Youtube 2160p 4K Ultra HD → Sicherstellen, dass Audio- und Videoexport markiert, sind → Bitrate 53-68 Mbps → Maximale Renderqualität markieren → Export.

Für den Video- und Audioexport in Adobe Premiere Pro wurden die Datenrate, das Dateiformat sowie die Video- und Audiocodierung entsprechend den Empfehlungen des

Video-Streaming-Diensts YouTube eingestellt, da die Videos anschließend über die entsprechende Streaming-Plattform veröffentlicht werden (Youtube-Hilfe, 2022). Das Seitenverhältnis für 360°-Videos beträgt 1:1 und ist somit quadratisch. Dies ist jedoch kein Hindernis, da sich der Youtube-Player automatisch an die Größe des Videos anpasst (YouTube-Hilfe, 2022).

6.2.4 Veröffentlichung

Sobald das 360°-Video bearbeitet und exportiert wurde, kann dies über YouTube veröffentlicht werden. Das Veröffentlichen von 360°-Videos auf YouTube ist genauso wie das Veröffentlichen von 2D-Videos. Auf der YouTube-Webseite muss man nur auf die kleine Kamera (vg. Abb. 46) oben rechts in der Nähe des Benachrichtigungssymbols klicken, dann „Video hochladen“ wählen, schließlich einfach die Videodatei, egal ob 360° oder 2D, in das entsprechende Fenster ziehen.

Abb. 46: Kamerasymbol auf der Youtube-Website (YouTube-Webseite, 2022)

Visualisierung von 360°-Videos mit VR-Brille

Um das 360 °-Video mit der HP Reverb 2 VR-Brille anschauen zu können, muss dieser via USB oder USB-C und DisplayPort mit dem Computer verbunden werden. Über den Edge-Browser in der Mixed-Reality-App kann der Nutzer unkompliziert YouTube durchsuchen und VR-Videos ansehen, indem er das VR-Brillen-Symbol im YouTube-Videoplayer auswählt, und somit wird das entsprechende Video in 360° auf dem Headset angezeigt.

7 Auswertung und Interpretation der Ergebnisse

Im folgenden Kapitel werden die im Rahmen der Studie erhobenen Daten vorgestellt, ausgewertet und interpretiert. Zu Beginn des Kapitels werden die und Unterschiede und Gemeinsamkeiten zum Einsatz des 360°-Aufnahmeformates in der Veranstaltungsbranche abgebildet. Darauf aufbauend sollen branchenübergreifen Schlussfolgerungen und Maßnahmen für den Einsatz im Veranstaltungstechnikbereich abgeleitet werden. Anschließen sollen neue Erkenntnisse dargestellt und erläutert werden.

Um eine Auswertung und spätere eindeutige Vergleichbarkeit der Videoaufnahmeformat gewährleisten zu können, wird die 5- stufigen Skala der Fragestellungen 5 und 6, in eine 3- stufige Skala umkodiert, die lediglich noch die Tendenzen: „besser, gleich, schlechter“ wiedergibt (Kirchhoff et al., 2010).

1= viel schlechter 2= schlechter	zu 1 = schlechter
3= gleich	zu 2 = gleich
4= besser 5= viel besser	zu 3= besser

Für die Auswertung des Gefühls der „Umhüllung“ und „vor Ort Sein“, wird die 5- stufigen Skala der Fragestellungen 7;8;9 und 10, ebenfalls in eine 3- stufige Skala nach Intensität umgewandelt. Die vorliegende Beurteilungsskala beinhaltet nun die Tendenzen: „nicht, neutral, intensiv“. Diese wird wie folgt dargestellt:

1= gar nicht 2= Umhüllung nicht besonders bemerkbar/ nicht besonders	zu 1 = nicht
3= neutral/ normal	zu 2 = neutral
4= etwas umhüllt/ etwas 5= komplett umhüllt/ sehr stark	zu 3= intensiv

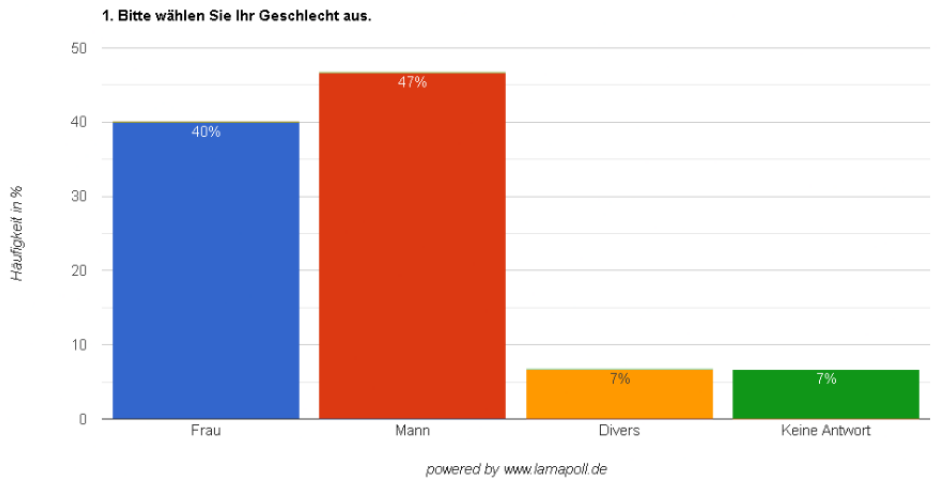
Für eine präzise Einschätzung der Kaufbereitschaft der befragten Probanden, soll im Folgenden die 7- stufigen Skala der Frage 12, in eine 3- stufige Skala mit den Tendenzen: „stimmt nicht zu“, „neutral“, „stimmt zu“ umkodiert werden. Wie bereits erwähnt, bezieht sich die Fragestellung auf die mögliche Akzeptanz von potentiellen Besuchern für 3D-VR-Veranstaltungen einen bestimmten Betrag zu bezahlen. Dazu wird die Beurteilungsskalen transformiert von:

1= -10€ oder weniger	zu 1 = stimmt nicht zu
2= -5€	
3= -1€	
4= 0€	zu 2 = neutral
5= 5€	zu 3= stimmt zu
6= 1€	
7= 10€ oder mehr	

Die Abstufung der Skala auf drei mögliche Ausprägungen, soll später für die Übersichtlichkeit in der Darstellung der Ergebnisse, und die Lesbarkeit der Tabellen, bei der Interpretation der Daten erleichtern (Kirchhoff et al., 2010).

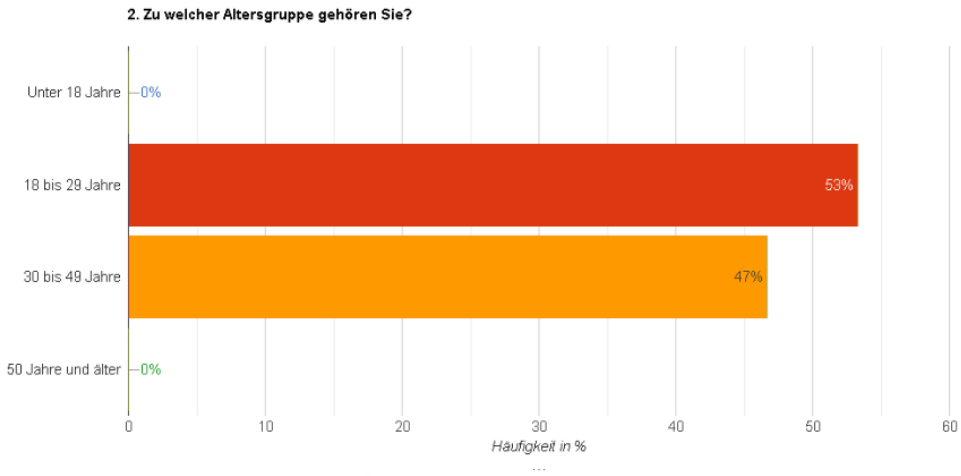
7.1 Allgemein

Frage 1.



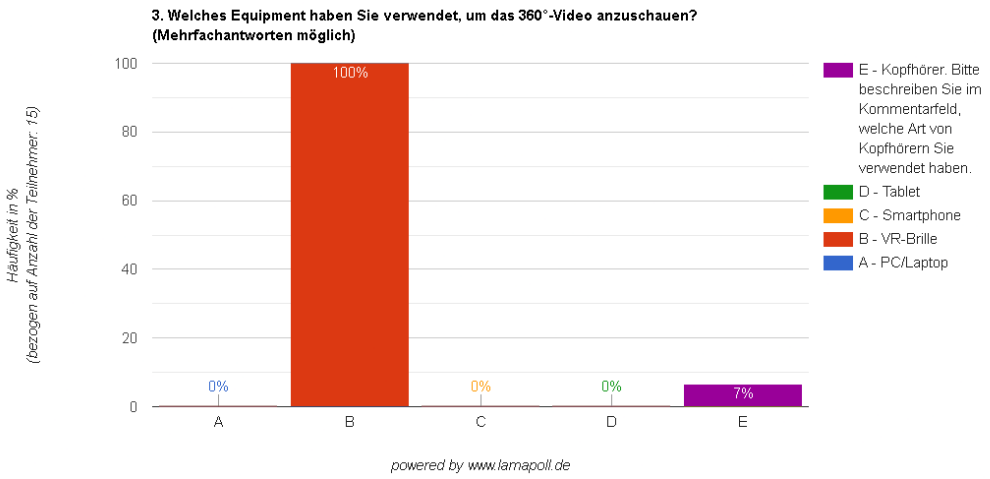
Von den 15 Teilnehmenden waren 7 Personen Männer, 6 Personen Frauen sowie 1 Person war Diverse. Weiterhin hat ein Proband die Frage mit „Keine Antwort“ beantwortet. Für diese Studie wurden Männer und Frauen in möglichst gleichem Maße zur Teilnahme eingeladen, um die Geschlechterrepräsentation so ausgewogen wie möglich zu halten und relevantere Ergebnisse zu erzielen. Alle Probanden haben die erste Frage beantwortet.

Frage 2.



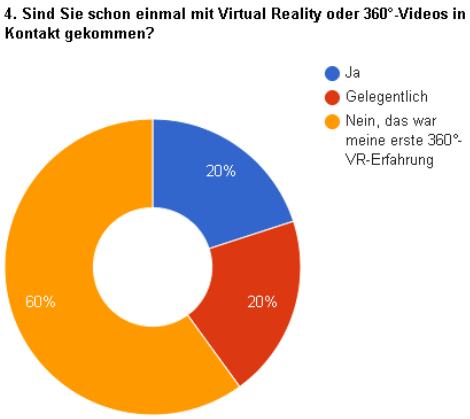
53 % der Teilnehmenden bzw. 8 Probanden gehören der Altersgruppe von 18 bis 29 Jahren an. Die restlichen 7 Probanden entsprachen der Altersgruppe von 30 bis 49 Jahren.

Frage 3.



Alle eingeladenen Teilnehmenden haben während des Versuches eine VR-Brille verwendet. Die VR-Brille ist mit eingebauten Lautsprechern ausgestattet, welche als Kopfhörer verwendet wurden, um einen immersiven Klangeindruck zu gewährleisten. Dieses wurde von einem der Teilnehmenden im Kommentarfeld wiederholt erwähnt.

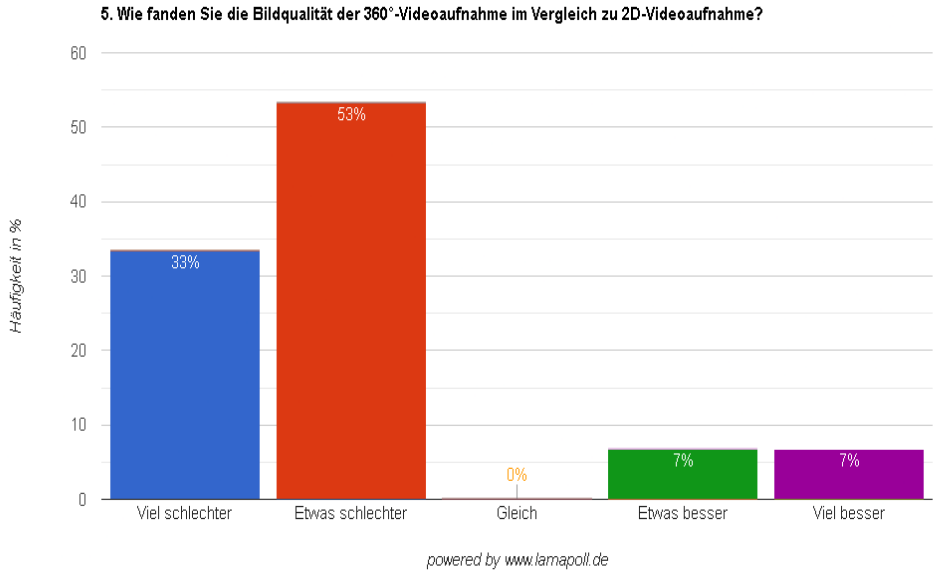
Frage 4.



60 % der Befragten bzw. 9 Probanden, haben noch nie ein 360°-Video mit einer VR-Brille angeschaut oder aber haben keine Erfahrung mit Virtual Reality. 20 % bzw. 3 Probanden haben bereits Erfahrungen mit VR-Technologie und 20 % nutzen diese gelegentlich. Die Mehrheit der Befragten ist demnach nicht mit Virtual Reality oder 360°-Videos in Berührung gekommen. Hieraus könnte geschlussfolgert werden, dass die VR-Technologien bzw. 360°-Aufnahmeformat im allgemeinen Konsumerbereich bis dahin nicht weitverbreitet sind.

7.2 Ergebnisse des Vergleichs des Sinneseindrucks Aspekt

Frage 5.

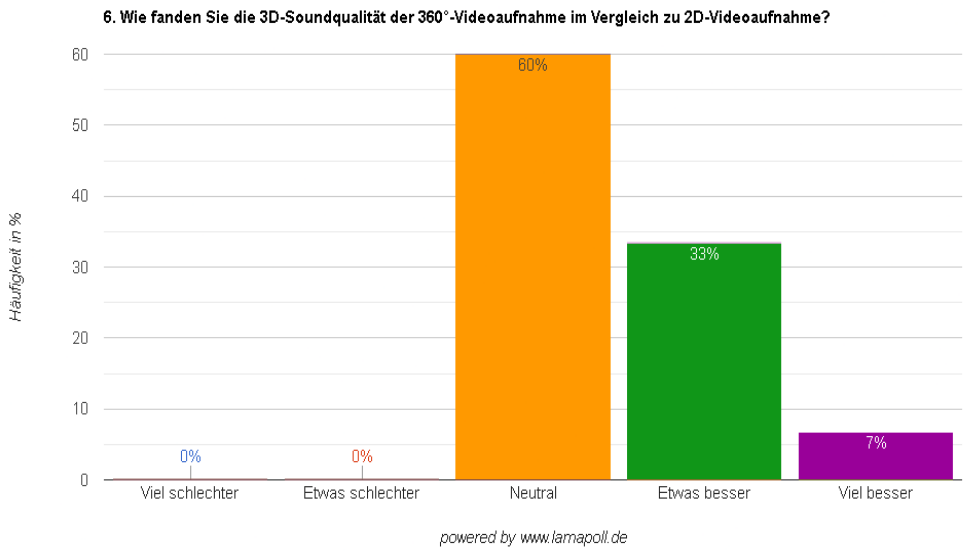


53% der Befragten bzw. 8 Probanden fanden die Bildqualität des 360°-Videos etwas schlechter als die des 2D-Videos. 33% bzw. 5 Probanden empfanden die Bildqualität viel schlechter, 7% bzw. 1 Personen etwas besser sowie die restlichen 7% viel besser. Insgesamt fanden 86% bzw. 13 Probanden die Bildqualität des 360°-Videos schlechter als die des 2D-Videos und 14% bzw. 2 Probanden fanden sie besser. Dieses lässt sich einerseits durch die Verwendung einer Verbraucher-VR-Kamera erklären, die keine ideale Bildauflösung bietet. Insbesondere wird dieses bei Aufnahmen im Dunkeln als auch durch das Fehlen von Fokus- und Zoom-Optionen erkennbar. Diese Optionen sind zur Anpassung der Schärfe des Bildes unverzichtbar.

Andererseits sollte festgehalten werden, dass im Gegensatz zu konventionellen Bildschirmen mit einer Auflösung von Wide XGA (1.360 x 769 Pixel) bis Full HD (1.920 x 1.080 Pixel), die nur etwa 30°- des horizontalen Sichtfeldes ausfüllen, verteilen sich die 2.160 x 1.200 Pixel aktueller hochauflösender stationärer VR-Brillen auf einen horizontalen Blickwinkel von 90°- bis 110°. D.h., dass die maximale Auflösung innerhalb desselben Sichtfeldes bei einer VR-Brille im Vergleich zu Computer-Anzeigen etwa um den Faktor 3 bis 4 reduziert wird (Hebbel-Seeger, 2018). Dementsprechend werden Menschen und Gegenstände ab einer Entfernung von ca. 7-10 m kaum noch

wahrgenommen (Hebbel-Seeger, 2018). Infolgedessen ist die Bildqualität bei einer 4K-360°-Videoaufnahme nur für Objekte ausreichend, die sich in unmittelbarer Nähe befinden (Windscheid und Gold, 2022). Für die 2D-Videoaufnahme wurde eine semiprofessionelle Kamera verwendet, die mit fast allen Tools und Optionen (Focus, Zoom, Weißabgleich usw.) ausgestattet ist, die dem Benutzer helfen, ein hochwertiges Bild in 4K zu erhalten. Das resultierende Bild wurde dann während des Experiments auf einem 4K-Computerbildschirm angezeigt.

Frage 6.

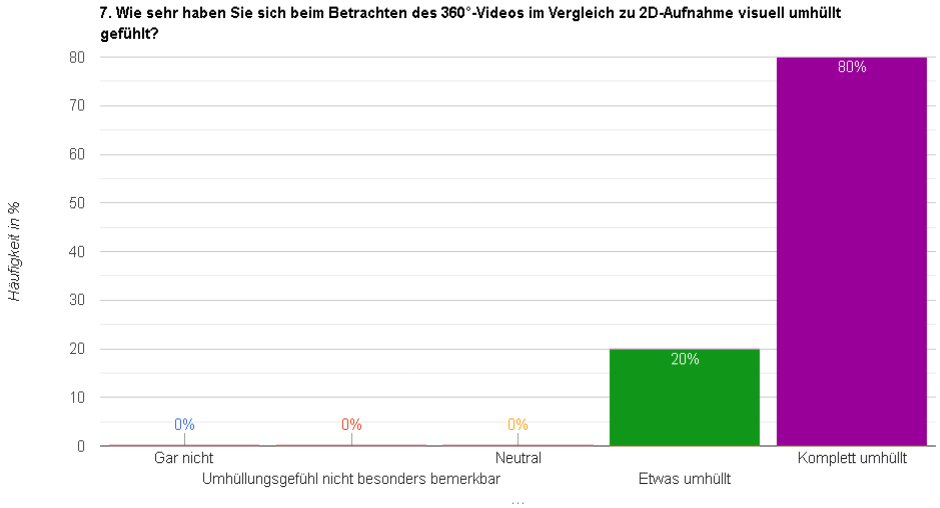


Bei der Auswertung der Frage 6 konnten 60% der Befragten bzw. 9 Probanden, keinen Unterschied zwischen dem 3D-Ton beim Betrachten des 360°-Videos gegenüber dem Stereoton der 2D-Aufnahme feststellen. 33% bzw. 5 Personen fanden die Qualität des 3D-Tons besser und 7% finden diese viel besser. Zusammengefasst konnten 60% keinen Unterschied zwischen den beiden Tonformaten feststellen (gleiche Tonqualität) und 40% bzw. 6 Personen hielten diese für besser.

Hier könnte die neutrale Tendenz verdeutlicht werden, denn der 3D-Sound unterscheidet sich zwar von Stereo und bietet ein breiteres Hörspektrum bzw. natürlicher Klang, aber hier wurde der Youtube-Webplayer verwendet, der zwar das Raumklangformat First Order Ambisonics (FOA) mit Head-Locked-Stereo unterstützt, dabei ändert sich aber der Ton nicht, wenn der Zuschauer sein Kopf bewegt (YouTube-Hilfe, 2022). Um eine natürlichere Hörumgebung zu schaffen und ein besseres Gefühl

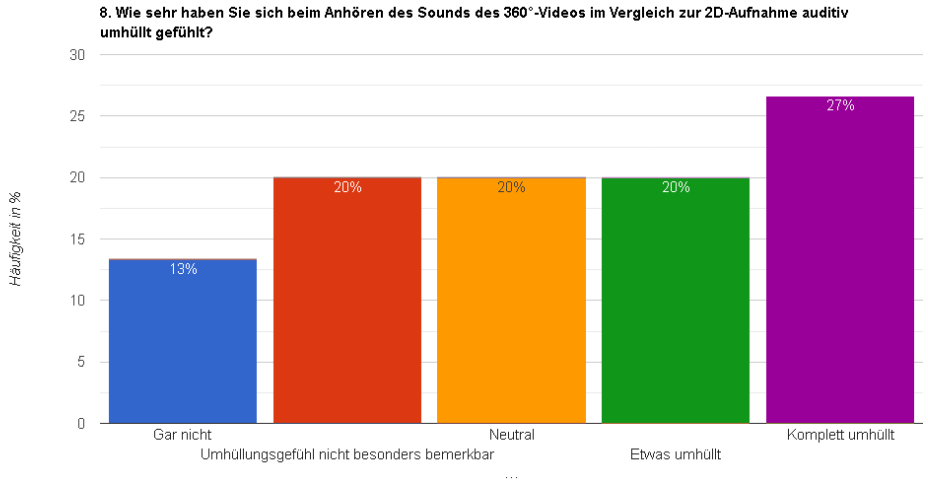
der Immersion und Präsenz zu erzeugen, muss es jedoch einen Ausgleich geben, wenn der Zuhörer seinen Kopf dreht, damit die virtuelle Klangquelle fest an ihrem Platz bleibt, wenn der Zuhörer seinen Kopf bewegt - so wie es in der natürlichen Hörumgebung der Fall ist (Roginska & Geluso, 2018). Dies könnte eine Erklärung dafür sein, warum sich hier eine eher neutrale Meinung herausgebildet hat. Denn das Ergebnis ist hier ein statisches Hörerlebnis (Roginska & Geluso, 2018).

Frage 7.



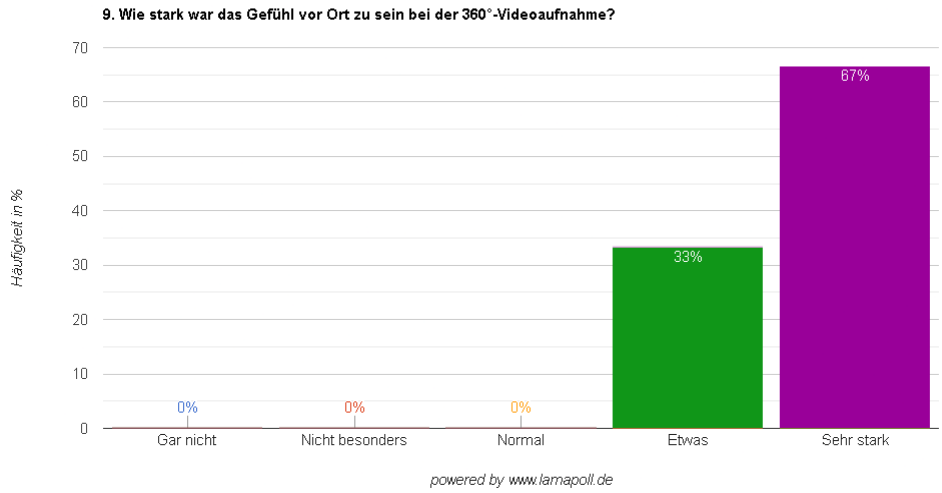
Anhand dieser Frage soll herausgefunden werden, wie hoch das Gefühl der visuellen Umhüllung in Relation zu konventionellen 2D Aufnahmen ist. Auffällig ist, dass sich 80 % der Befragten, bzw. 12 Probanden, „komplett umhüllt“ fühlten. Die restlichen 20% fühlten sich während der Untersuchung nur „etwas umhüllt“. Hier war das Gefühl der Umhüllung oder Immersion, durch die Verwendung der stationären VR-Brille von allen Probanden sehr intensiv bemerkbar. Im Gegensatz zu 2D-Videoformaten bieten 360°-Videos eine Rundum-Perspektive, die bei allen Probanden ein vollständiges Umhüllungsgefühl erzeugt.

Frage 8.



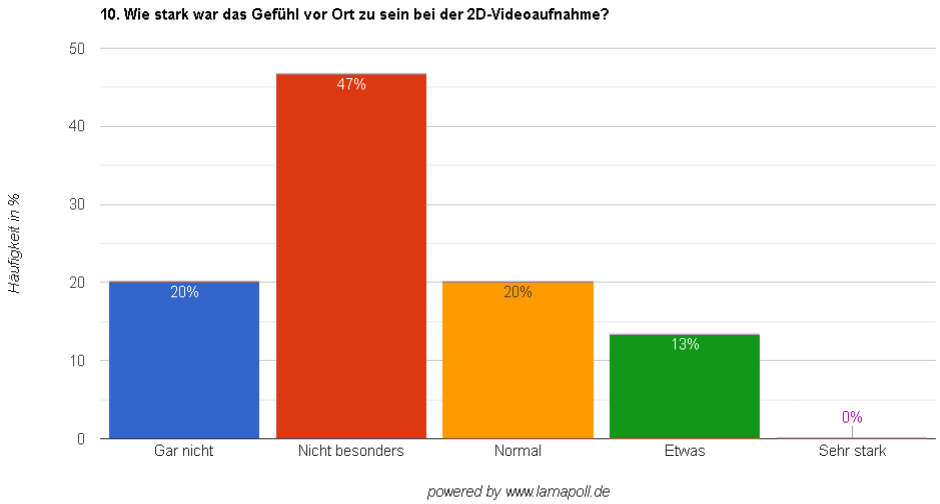
Bei der Bewertung des Gefühls von der 3D-Tonumhüllung im Vergleich zum Stereoton der konventionellen Aufnahme fühlten sich 27% der Befragten bzw. 4 Probanden komplett umhüllt, 20% bzw. 3 Probanden haben sich etwas umhüllt gefühlt, 20% waren neutral, für 20% war die 3D-Tonumhüllung nicht besonders spürbar und 13% bzw. 2 Probanden bemerkten keinen Unterschied. Zusammengefasst empfanden 47% bzw. 7 Probanden die 3D-Klangumhüllung intensiv, 20% jeweils neutral, und 33% bzw. 5 Probanden haben die 3D- Tonumhüllung erst gar nicht bemerkt. Hier ist der Unterschied nicht spürbar, da der 3D-Ton sich kaum spürbar von der Stereoaufnahme unterscheidet. Die Gründe hierfür wurden bereits bei der Auswertung der Frage 6 genannt.

Frage 9.



67% der Befragten bzw. 10 der Teilnehmenden fühlten sich beim Betrachten der 360° - Aufnahmen so, als wären sie bei der Veranstaltung „vor Ort zu sein“. Die restlichen 5 Probanden gaben an, sich teilweise bzw. etwas wie „vor Ort zu“ zu fühlen. Demnach hatten alle Teilnehmenden die Illusion, während der Veranstaltung im Zuschauerbereich zu sitzen.

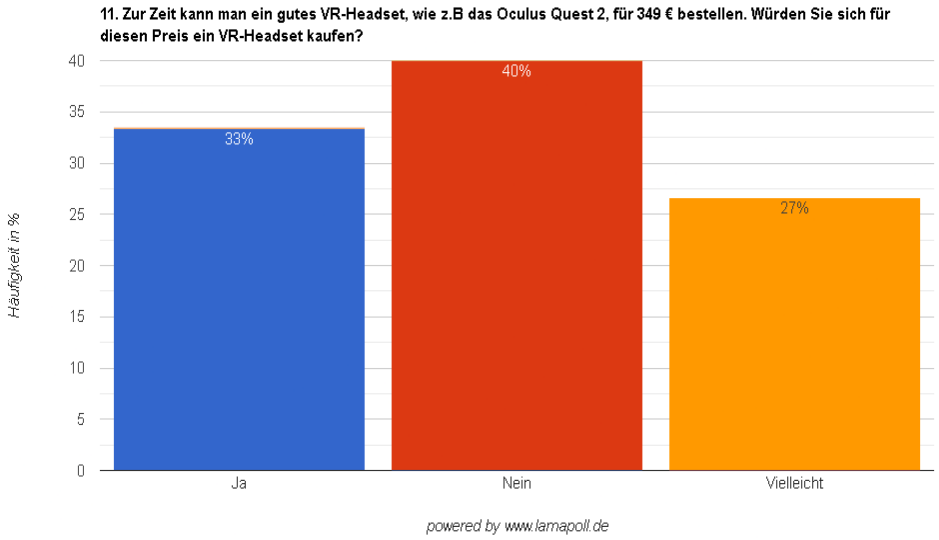
Frage 10.



Beim Betrachten der 2D-Videoaufnahme, fühlten sich 47% der Befragten bzw. 7 Probanden nicht besonders vor Ort, 20% oder 3 Personen fühlten sich gar nicht vor Ort, 20% haben eine neutrale Meinung und 13% bzw. 2 Probanden fühlten etwas vor Ort. Zusammenfassend lässt sich sagen, dass das Gefühl des vor Ort seins, bei der 2D-Aufnahme von 67% bzw. 10 Probanden nicht empfunden wurde, 20% sind neutral, 13% haben etwas Intensives gespürt. Insbesondere im Vergleich zu dem Ergebnis der Frage 9 wird hier ersichtlich, dass nur ein geringer Teil der Probanden das Gefühl hat, beim Betrachten der 2D- Aufnahmen wirklich „teil der Show“ zu sein. Dies lässt sich dadurch erklären, dass das 2D-Videoformat das Sichtfeld des Betrachters nicht vollständig abdeckt, was zu einer geringeren Umhüllung oder Immersion führt, was wiederum ein viel geringeres Gefühl der Involviertheit und des vor Ort seins Gefühls bzw. Präsenz erzeugt.

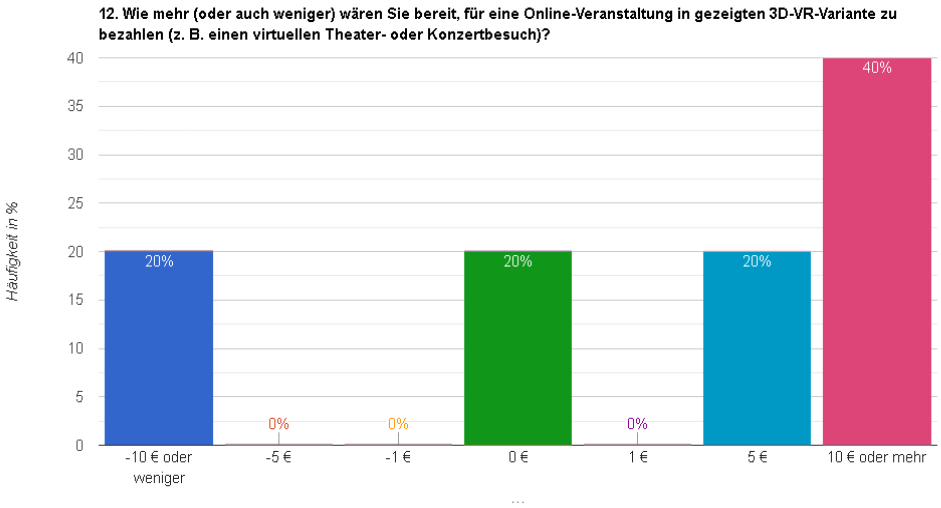
7.3 Ergebnisse der Bewertung des wirtschaftlichen Aspektes

Frage 11.



40% der Befragten bzw. 6 Personen finden den Preis der VR-Brille zu teuer und sind nicht bereit, so viel Geld dafür auszugeben. Weitere 33% bzw. 5 Personen wären bereit den angegebenen Preis dafür zu bezahlen und 27% bzw. 4 Personen sind unentschlossen. Hieraus kann zum einen geschlussfolgert werden, dass die aktuellen Markpreise für VR-Brillen noch zu hoch sind oder aber vielleicht auch, prinzipiell kein Interesse besteht, diese Form des Entertainments zu nutzen.

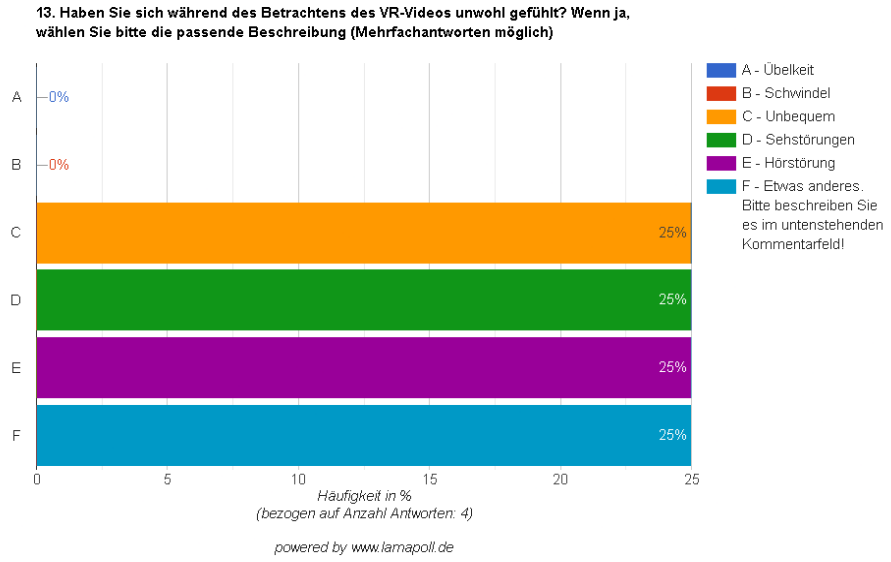
Frage 12.



Mit dieser Frage soll untersucht werden, wie hoch die Bereitschaft ist, sich ein Ticket für eine 3D-VR Online-Veranstaltung zu kaufen. Hierbei antworteten 40% der Befragten, dass diese bereit wären, mehr als 10€ und 20% der Teilnehmenden jeweils 5€ für eine VR-Veranstaltung auszugeben. Eine relativ neutrale Meinung haben 20% der Befragten. Auch würden 20% tendenziell 10€ und weniger für ein Ticket dieser Art ausgeben. Zusammenfassend kann festgehalten werden, dass 60% der Befragten bereit sind, mehr Geld für Online-Veranstaltungen im 360° gezeigten Format auszugeben. Demnach kann anhand der Ergebnisse geschlussfolgert werden, dass der Einsatz von 360°-Videoaufnahmen im Veranstaltungsbereich im Vergleich zu konventionellen Videoaufnahmen für den allgemeinen Verbraucher durchaus sinnvoll erscheint.

7.4 Ergebnisse der Bewertung der Nutzerkomfort Aspekt

Frage 13.



Die Frage 13 untersucht mögliche Nebenwirkungen bei der Verwendung von VR-Brillen. Beantwortet haben diese Frage 4 Teilnehmende (jeweils 2x männlich und 2x weiblich), wobei alle untersuchten Altersgruppen hierbei vertreten sind. Eine weibliche Person gab an, visuelle Störungen beim Betrachten des 360°-Videos zu empfinden. Weitere männliche Personen verspürten zum einen Hörstörungen während der Vorführung, sowie eine andere klagte darüber, dass die VR-Brille unbequem sei.

Aus diesem Sachverhalt können allerdings keine präzisen Schlussfolgerungen gezogen werden, da die Ergebnisse keine Zusammenhänge zwischen dem Geschlecht, als auch der Altersgruppe und VR-Erkrankungen aufzeigen. Da 27 Prozent der Probanden diese Frage beantworteten, kann dementsprechend davon ausgegangen werden, dass 73 Prozent der Probanden während des Experiments kein Unwohlsein empfunden haben. Mögliche Ursachen für das Unwohlsein einzelner Personen könnten das Tragen eines ungewöhnlich großen Geräts auf dem Kopf sein. Die minderwertige Bildqualität stellt jedoch die Hauptursache dar. Die VR- Erkrankungen zeigen sich in keiner hohen Ausprägung. Es ist davon auszugehen, dass die kurze Filmdauer, circa 6 Minuten und das Stillsitzen ohne simulierte Bewegungen das Unwohlsein minimieren. Bei der Hörstörung wird die Lautstärke in diesem Fall als Hauptursache angenommen.

Frage 14.

Antworten zu Anmerkung
Die Bildqualität war 3D etwas schlechter als in 2D weswegen es sich eher unwirklich angefühlt hat.
Das war ein interessantes Erlebnis, die 3D video Qualität war etwa verschwommen.
Sehr schöne immersive Erfahrung
Das ist sehr besondere neue Erfahrung :)

Von 15 Teilnehmenden haben 4 die Frage *„Bitte kommentieren Sie hier, was Ihnen im Zusammenhang mit Ihren abgegebenen Bewertungen evtl. aufgefallen ist/was Sie uns näher erläutern bzw. mitteilen wollen:“* beantwortet, was 27 % der Gesamtzahl der Befragten entspricht. Hier konnten keine neuen Informationen aus den Antworten gewonnen werden. Allerdings kritisierten zwei Personen die Bildqualität der 360°-Aufnahme, eine Person hält das Ziel des Erlebnisses aufgrund der schlechten Bildqualität für nicht erfüllt, die anderen 3 Teilnehmenden halten das Erlebnis jedoch für besonders und betiteln dies als interessante und neue Erfahrung. Mögliche Ursachen für eine schlechte Bildqualität der 360°-Videoaufnahmen wurden bereits in der Frage 5 genauer erläutert.

Beurteilung der untersuchten Vergleich Aspekte

In diesem Kapitel wird ausgehend von den Ergebnissen der erhobenen Daten, im allgemeinen Teil ein Gesamtüberblick über das Profil der Teilnehmenden, die an diesem Experiment teilgenommen haben, vorgestellt. Zunächst werden die Ergebnisse zusammengefasst und die drei Vergleichsaspekte einzeln anhand von Spitzenwerten bewertet, um zu ermitteln, ob diese Vergleichsaspekte den Einsatz von 360°-Aufnahmeformaten in der Veranstaltungsbranche benachteiligen oder ob sie ein überzeugendes Argument für die Nutzung dieser Technologie in diesem Bereich darstellen.

Zusammenfassung

An dieser Studie nahmen 15 Probanden mit gleichmäßiger Geschlechtsrepräsentation, im Alter von 18 bis 49 Jahren teil. Während der Studie benutzten alle eine stationäre VR-Brille, die HP Reverb G2 Virtual-Reality-Headset, die mit eigenen Lautsprechern ausgestattet ist und einen immersiven dreidimensionalen Raumklang wiedergibt (HP-Website, 2021). 60% der Befragten sind noch nicht mit der Virtual Reality oder 360°-Videos in Berührung gekommen. VR bzw. 360°-Videoformat ist also für die Mehrheit der Befragten nicht bekannt.

Beim Vergleich der Bildqualität fanden 86% der Teilnehmenden die Bildqualität der 360°-Videoaufnahme schlechter als die der 2D-Videoaufnahme. Eine 4K-360°-Kamera ist daher für die Aufnahme ferner Objekte ungeeignet. Alle Befragten beschrieben das Gefühl, umhüllt und vor Ort zu sein, im Gegensatz zum 2D-Video, bei welchem dieses Gefühl gar nicht zu spüren war. Die 360°-Videoaufnahme löst psychovisuelle Empfindungen aus, die bei 2D-Aufnahmen nicht wahrgenommen werden können.

Die Tendenz beim Vergleich von 3D- und Stereo-Klangqualität ist eher neutral, da die Teilnehmenden meist keinen Unterschied feststellen konnten. Dieses gilt ebenso im Bezug zur Ausprägung einer möglichen Umhüllung des 3D-Klangs in Relation zu einem konventionellen Stereosignal, welches von den Befragten ebenfalls als nicht stark wahrgenommen wurde. Der Youtube-Webplayer ermöglicht die Wiedergabe des 3D-Ambisonic-Tons, scheint aber nicht geeignet zu sein, um die voreingestellten kopfbezogenen Signale zu verfolgen.

Bei der Messung der Kaufbereitschaft ist die Meinung der Befragten gemischt, wobei eine signifikante Anzahl von Personen, nämlich 40 %, den Kauf einer VR-Brille verneinte. Allerdings waren 60% bereit, mehr Geld für eine Online-Veranstaltung in der gezeigten 3D-VR-Variante auszugeben. VR-Erlebnisse werden daher von den Verbrauchern generell wertgeschätzt, aber der Zugang zu dieser Technologie ist immer noch zu teuer. Zur Bewertung des Aspekts der Bequemlichkeit bzw. des des Aspekts des Nutzerkomforts kann davon ausgegangen werden, dass das Experiment für den überwiegenden Teil (80 %) der Befragten als angenehm empfunden wurde. Um auch hier ein repräsentativeres Meinungsbild wiedergeben zu können, sollte weitere Versuche mit längeren Videoaufnehmen durchgeführt werden.

Weitere Nachteile und Vorteile der Nutzung sind übersichtlich in der Folgende Tabelle dargestellt:

Nachteile	**Vorteile**
● Die 4K-Bildqualität der 360°-Aufnahme wird als qualitativ schlechter wahrgenommen als 4K-2D-Aufnahme. ● Bislang aufwändig und teuer für den durchschnittlichen Verbraucher. ● 3D-Audio wird von aktuellen Web-Playern nicht richtig unterstützt im Gegensatz zu Stereo. ● VR-Krankheiten können möglicherweise bei bewegten Videobildern und langer Videodauer auftreten.	● Erwecken von neuen psychovisuellen Sinneseindrücken, die sonst nur in einer Live-Show zu spüren sind, wie das Gefühl des Vor-Ort-Seins und des Umhüllungsgefühls. ● Das 360°-Videoformat wird vom virtuellen Betrachter gegenüber dem 2D-Videoformat bevorzugt, wenn es sich um ein Konzert oder eine Theateraufführung handelt. ● Die Nutzung einer VR-Brille ist bequem. (Bedingt durch eine statische Aufnahme und ein kurzes Video).

8 Fazit und Ausblick

Im Rahmen dieser Masterarbeit wurde von einer Veranstaltung eine 360°-Videoaufnahme mit 3D-Klang aufgezeichnet und das produzierte Ergebnis, mit einer zeitgleich erstellten konventionellen 2D-Videoaufnahme durch ausgewählte Probanden bewertet. Es wurde ein Fragenkatalog mit 13 Fragen in den Bereiche Sinneswahrnehmung, Wirtschaftlichkeit und Nutzerkomfort entwickelt. 15 Probanden nahmen an dem Experiment teil und anhand der Ergebnisse konnte ein Meinungsbild erstellt werden.

Die Ergebnisse der vorliegenden Studie belegen, dass die verwendete Methode neue psychovisuelle Eindrücke für die Zuschauer erzeugen kann. Das Gefühl der Umhüllung bzw. Immersion und das Gefühl des „vor Ort seins“, wurden von allen Teilnehmenden der Untersuchung als sehr stark empfunden. Diese Gefühle stehen im Zusammenhang mit der visuellen Wahrnehmung und sind sonst nur in echten Live-Veranstaltungen spürbar. Folglich wird ersichtlich, dass VR-Inhalte oder 360°-Videos durch den Perspektivwechsel und dem Rundum-Sichtfeld eine Erweiterung des Wahrnehmungshorizonts für Zuschauer:innen ermöglichen können. Es kann nicht nur ein höheres Interesse und Spannung für die Zuschauer:innen erzeugt werden, sondern haben bspw. Schauspieler:innen hierdurch einen größeren Spielraum für künstlerisches Arbeiten oder Bühnenbildner:innen mehr Gestaltungsmöglichkeiten für die Konzipierung von Bühnenbildern.

Auch musste festgestellt werden, dass die Bildqualität der verwendeten VR-Kamera für die meisten Teilnehmenden als nicht zufriedenstellend empfunden wurde. Demnach sollten hier Maßnahmen zur Verbesserung eingeleitet werden. Eine Möglichkeit wäre, die Verwendung mehreren 4K-VR-Kameras, die in jeweils unterschiedlichen Entfernungen vom jeweiligen Aufnahmefokus aufgestellt werden. Hierdurch kann der Betrachter zum einen aus verschiedenen Blickwinkeln und Positionen selbstständig wählen und zum anderen verbessert sich durch den verringerten Abstand auch die Qualität der Videoaufnahme. Eine weitere Lösungsmöglichkeit wäre der Einsatz von VR-Kameras mit besserer Auflösung. Leider sind Kameras mit hoher Bildqualität im

Konsumerberreich aktuell noch sehr teuer. Sollen professionelle Aufnahmen von Veranstaltungen erstellt und hierfür Tickets an Kunden verkauft werden, sollte für eine Professionelle VR-Kamera mit guter Bildqualität ausreichend investiert werden. Der 3D-Sound war für die Teilnehmenden kaum wahrnehmbar, da die meisten aktuellen Streaming-Anbieter keinen echten, räumlich getrackte 3D-Ambisonic-Sound wiedergeben können. Als einzige Ausnahme ist hier Facebook zu nennen, welches allerdings aus Datenschutzgründen in dieser Untersuchung nicht verwendet werden konnte. Um die Qualität der 3D-Audioaufnahmen zu erhöhen, ist es aber prinzipiell sinnvoller, neben dem 3D-Hauptmikrofon zusätzliche Mikrofone für die Aufnahmen im Veranstaltungsort aufzustellen. Mit anschließender Nachbearbeitung durch eine professionelle Audio-Lokalisierungssoftware lassen sich hierdurch sehr gut Ergebnisse erzielen (Winscheid und Gold 2022). Weiterhin wurde festgestellt, dass Verbraucher:innen tendenziell bereit wären für eine VR-Veranstaltung mehr Geld auszugeben. Allerdings wird der Großteil der verfügbaren Hardware für diese als noch zu teuer empfunden. Hierbei muss in Zukunft beobachtet werden, ob die jeweiligen Hardware kostengünstiger im Verkauf, sich die Akzeptanz der Nutzer:innen hierfür mehr Geld auszugeben verändert. Weiterhin muss an dieser Stelle festgehalten werden, dass aufgrund der geringen Probandenzahl keine repräsentativen Ergebnisse in dieser Forschungsarbeit ermittelt werden konnten. Insbesondere im Bereich des Nutzerkomforts werden weitere Befragungen benötigt, um qualitativ hochwertige Ergebnisse erzielen zu können. Daher sollte die vorliegende Studie zum Aufbau weiterer Studien genutzt und die einzelnen Forschungsergebnisse weiter vertieft werden.

Abschließend kann festgehalten werden, dass die Verwendung der 360°-Aufnahme im Bereich der Veranstaltungstechnik einen bereichernden und inspirierenden Beitrag zur kreativen Erweiterung einer Performance geben kann. Ebenfalls können online Übertragungen erfahrbarer und realistischer erlebt werden als konventionelle 2D-Aufnahmen. Die gleichzeitig ortsungebundene Teilnahme und dennoch das Gefühl des Vor-Ort-Seins bieten unterschiedlichen Personenkreisen die Möglichkeit Veranstaltungen zu besuchen, ohne dafür das Haus verlassen zu müssen. Somit haben die Künstler:innen die Chance trotz möglicher Pandemie o.ä. die Veranstaltung für die Zuschauer:innen zu einem unvergesslichen Erlebnis zu machen.

Abbildungsverzeichnis

Abb. 1: Darstellung einer Anzeige, die die Zwischenspeicher für ein 30-minütiges Streaming-Video mit beschrifteten Segmenten und Punkten (Comer, 2014) 3

Abb. 2: Empfohlene kontinuierliche Geschwindigkeit (YouTube-Hilfe, 2022)..... 7

Abb. 3: Querschnitt des menschlichen Auges (Schmidt, 2012)..... 13

Abb. 4: Bau der Netzhaut, Zapfen und Stäbchen (Klett, 2014) 13

Abb. 5: Aktuelle Videostandards (Bühler et al., 2018) 14

Abb. 6: Beispiel Seitenverhältnis (eigene Darstellung)..... 14

Abb. 7: Additive Farbmischung Drei Lichtpunkte der der

Primärfarben RGB (Böhringer et al., 2014) 15

Abb. 8: Beispiel der Umwandung von RGB zuYCrCb (Wikipedia, 2012) 16

Abb. 9: Ivan Sutherlands Head-Mounted Display (Jewst, 2021) 20

Abb. 10: Stationäre HP Reverb G2 VR-Brille (HP-Webseite)..... 22

Abb. 11: Mobile Google Cardboard VR-Brille (Feyder und Rath-Wiggins, 2018)..... 22

Abb. 12: Links: Mit dem „3DoF“ Tracking-System, kann sich der Benutzer von einem festen Punkt aus in der virtuellen Welt umschauen. Rechts: Mit dem „6DoF“ Tracking-System kann sich der Benutzer in der virtuellen Welt wie in der realen Welt bewegen (Magnor, Sorkine-Hornung, 2020) 23

Abb. 13: Aufnahmewinkel einer 360°-Kamera. Links: zwei Kameralinsen, Rechts: vier Kameralinsen (Windscheid und Gold, 2022) 24

Abb. 14: Vuze Kamera ausgestattet mit 8 Kameralinsen (Amazon, 2022) 24

Abb. 15: Physiologisches Modell des stereoskopischen Sehens. IPD, Interpupillarer Abstand (Burdea und Coiffet, 2003)..... 25

Abb. 16: Aufbau des menschlichen Ohrs (Bühler et al., 2018)..... 31

Abb. 17: Beispiel ein tetraedrisches Mikrofonsystem (Paterson und Lee, 2021) 34

Abb. 18: Tetraedrische Nummerierung der Mikrofonausgangskanäle (Paterson und Lee, 2021) 34

Abb. 19: Zoom H3-VR (Zoom -Webseite) 35

Abb. 20: Sennheiser AMBEO VR Mic (Sennheiser -Webseite) 35

Abb. 21: Røde NT-SF1 (Røde Webseite) 35

Abb. 22: Ambisonic B-Format-Komponenten; übereinandergelegt (ganz links) und separat (Mitte bis rechts) W, X, Y und Z dargestellt (Paterson und Lee, 2021) 36

Abb. 23: Die SN3D normalisierten sphärischen harmonischen Funktionen, welche in Ambisonics bis zur dritten Ordnung verwendet werden, gekennzeichnet mit alphabetischer (FuMa) und ACN-Notation (Paterson, Lee, 2021) 36

Abb. 24: Dreidimensionale Kopfrotation: Drehen (engl. yaw), Nicken (engl. pitch) und Schwenken (engl. roll) (Roginska und Geluso, 2018) 39

Abb. 25: Intensität der VR-Wahrnehmung (Feyder und Rath-Wiggins, 2018) 45

Abb. 26: Aufnahmeeinstellungen Sony PXW-X70 Kamera 53

Abb. 27: Eigenschaften Vuze 3D-360-Kamera (Humaneyes Technologies, 2017) 54

Abb. 28: Statusanzeigen Vuze 3D-360-Kamera (Humaneyes Technologies, 2017) 54

Abb. 29: Aufnahmeeinstellungen Vuze 3D-360-Kamera (Bildschirmfoto, Vuze Camera app) 55

Abb. 30: Einstellung des Ambisonic-Formats in ZOOM F6 (ZOOM CORPORATION, 2022) 56

Abb. 31: Kamerapotionen im Zuschauerbereich während der Echtzeit-Aufnahme im Bluemax Theater Berlin, (1) Vuze 3D 360 4K VR Kamera und Sennheiser AMBEO VR Mikrofon, (2) Sony PXW-X70 Kamera (Sghaier, 2022) 61

Abb. 32: Vuze 3D 360 4K VR Kamera, Sennheiser AMBEO VR Mikrofon und Zoom F6 Rekorder im Bluemax Theater Berlin (Sghaier, 2022) 63

Abb. 33: Sony PXW-X70 Kamera im Bluemax Theater Berlin (Sghaier, 2022) 63

Abb. 34: Parallelaufnahme Übersicht der Blue Man Group Show 2021 im Bluemax Theater Berlin, links ist die 2D-Videoaufnahme und rechts die 360°-Videoaufnahme zu sehen (YouTube, 2022) 63

Abb. 35: Fernsteuerung des Zoom F6 Rekoder durch die iOS-App F6 Control (Bildschirmfoto, iPhone 6) 64

Abb. 36: Output-Settings Übersicht in Vuze VR Studio Software, 360°-Video der Blue Man Group Show 2021 im Bluemax Theate Berlin (Vuze VR Studio, 2020) 66

Abb. 37: 360°-Video der Blue Man Group Show 2021 im Bluemaxtheater Berlin vor der Verwendung Refine Stitching Option (Vuze VR Studio, 2020) 66

Abb. 38: Verwendung der Refine Stitching Option in Vuze VR Studio Software, 360°-Video der Blue Man Group Show 2021 im Bluemaxtheater Berlin 66

Abb.39: Eigenschaften der 360°-Videodatei (Vuze VR Studio, 2020) 67

Abb. 40: 360°-Videomontage der Blue Man Group Show mit Adobe Premiere Pro (Adobe Premiere Pro, 2022) 68

Abb.41: AMBEO A-B-Konverter in Reaper 69

Abb. 42: EnergyVisualize in Reaper 69

Abb. 43: Audio 360 Meter Plugin in Reaper 70

Abb. 44: Audio 360 Meter Plugin in Reaper 70

Abb.45: Export der 3D-Audiodatei in Reaper 71

Abb. 46: Kamerasymbol auf der Youtube-Website (YouTube-Webseite, 2022) 73

Literaturverzeichnis

Adobe. (2022). Abgerufen am 24. August 2022 von Adobe Premiere Pro: https://www.adobe.com/de/products/premiere.html?mv=search&mv=search&sdid=LQLZT7BT&ef_id=a2b6735276fd177029251056b471eb96:G:s&s_kwcid=AL!3085!10!78821331035909!78821436163615

AlarmstufeRot. (2020). Abgerufen am 24. August 2022 von https://alarmstuferot.org/fakten

Amazon. (2022). Abgerufen am 24. August 2022 von Vuze - 3D 360° 4K VR Kamera - Schwarz: https://www.amazon.com/Vuze-3D-360°-Camera-Black/dp/B00K18V5VY?th=1

Amazonaws. (2017). Abgerufen am 24. August 2022 von Vuze Camera User Manual: https://s3.amazonaws.com/downloads.vuze.camera/manual/Vuze+Camera+User+Manual.pdf#

Böhringer, J., Bühler, P., Schlaich, P., & Dominik , S. (2014). *Kompendium der Mediengestaltung.* Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-642-54585-6

Brühwiler, H. (2022). *Masterarbeit: Binaurales Echtzeitabhörsystem an der Oper Graz: Methode zur Optimierung der Abhörsituation der Tonschaffenden.* Graz: FH JOANNEUM Graz, Institut für Elektronische Musik und Akustik Kunstuniversität Graz.

Bühler, P., Schlaich, P., & Sinner, D. (2018). *AV-Medien_Filmgestaltung – Audiotechnik – Videotechnik.* Berlin: Springer-Verlag. doi:https://doi.org/10.1007/978-3-662-54605-5

Burdea, G., & Coiffet, P. (2003). *Virtual Reality Technology.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Comer, D. (2014). *Internetworking with TCP/IP.* Edinburgh : Pearson Education Limited.

Demetriades, G. (2003). *Streaming Media Building and Implementing a Complete Streaming System.* Indianapolis, Indiana: Wiley Publishing, Inc.

Doerner, R., Broll, W., Grimm, P., & Jung, B. (2022). *Virtual and Augmented Reality (VR/AR): Foundations and Methods of Extended Realities (XR).* Springer Nature Switzerland AG. doi:https://doi.org/10.1007/978-3-030-79062-2

Dörner, R., Broll, W., Grimm, P., & Jung, B. (2019). *Virtual und Augmented Reality (VR/AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität.* Springer-Verlag GmbH Deutschland. doi:https://doi.org/10.1007/978-3-662-58861-1

Feyder, M., & Rath-Wiggins, L. (2018). *VR-Journalismus: Ein Handbuch für die journalistische Ausbildung und Praxis.* Springer Fachmedien Wiesbaden GmbH. doi:https://doi.org/10.1007/978-3-658-22217-8

Fischer, W. (2020). *Digital Video and Audio Broadcasting Technology A Practical Engineering Guide.* Munich: Springer Nature Switzerland AG. doi:https://doi.org/10.1007/978-3-030-32185-7

Furht, B. (2008). *Encyclopedia of Multimedia.* New York: Springer Science+Buisiness Media, LLC.

Github. (2022). Abgerufen am 3. Januar 2022 von SPATIAL WORKSTATION USER GUIDE: https://facebookincubator.github.io/facebook-360-spatial-workstation/

Google Play. (2022). Abgerufen am 24. August 2022 von Vuze Camera app: https://play.google.com/store/apps/details?id=camera.vuze&gl=US

Grand View Research. (2020). Abgerufen am 24. August 2022 von Virtual Reality In Gaming Market Size Worth $92.31 Billion By 2027: https://www.grandviewresearch.com/press-release/global-virtual-reality-in-gaming-market

Hebbel-Seeger , A. (2018). 360°-Video in Trainings- und Lernprozessen. In U. Dittler, & C. Kreidl, *Hochschule der Zukunft: Beiträge zur zukunftsorientierten Gestaltung von Hochschulen* (S. 265-290). Springer Fachmedien Wiesbaden GmbH. doi:https://doi.org/10.1007/978-3-658-20403-7

Hebbel-Seeger, A., Reiners, T., & Schäffer, D. (2014). *Synthetic Worlds: Emerging Technologies in Education and Economics.* Springer Science+Business Media New York. doi:DOI 10.1007/978-1-4614-6286-6

Heber, A. (2020). *Production Partner: Fachmagazin für Veranstaltungstechnik.* Abgerufen am 24. August 2022 von GRUNDLAGEN (LIVE-)STREAMING: Alles Streaming oder was?: https://www.production-partner.de/basics/alles-streaming-oder-was/

Hillmann, C. (2019). *Unreal for Mobile and Standalone VR Create Professional VR Apps Without Coding.* Singapore, Singapore: Springer Science+Business Media New York. doi:https://doi.org/10.1007/978-1-4842-4360-2

Hoffbauer, E. (2021). *MASTER'S THESIS: Evaluation of Surround Sound Setups based on Ambisonic Room Impulse Response Measurements.* Graz: University of Music and Performing Arts, Graz, University of Technology, Graz.

Hollenberg, S. (2016). *Fragebögen: Fundierte Konstruktion, sachgerechte Anwendung und aussagekräftige.* Springer Fachmedien Wiesbaden. doi:DOI 10.1007/978-3-658-12967-5

HP. (2022). Abgerufen am 24. August 2022 von HP Reverb G2: https://www.hp.com/de-de/vr/reverb-g2-vr-headset.html

Immergutrocken. (2021). Abgerufen am 24. August 2022 von Immergut Festival: https://immergutrocken.de/

Jewst, A. (2021). *Bachelorarbeit: Konzeption von Einsatzszenarien fur Mixed Reality-Technologie in der Grundlagenlehre der Elektrotechnik.* Hannover: Gottfried Wilhelm Leibniz Universität Hannover.

Kapser, P. (2021). *Bachelorarbeit: Entwicklung eines Lehr- & Laborprojekts zum Thema "AV-Streaming-Veranstaltungen" Hard- und Softwarelösungen im Vergleich.* Berlin: Beuth Hochschule für Technik Berlin. Von https://prof.bht-berlin.de/fileadmin/prof/alindau/Kapser_BEng.pdf abgerufen

Kirchhoff, S., Kuhnt, S., Lipp, P., & Schlawin, S. (2010). *Der Fragebogen: Datenbasis, Konstruktion und Auswertung.* VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden GmbH.

Klett. (2014). Abgerufen am 24. August 2022 von https://static.klett.de/software/shockwave/prisma_bio_ol/pb_tb02ov502/tb02ov502.html

Knoll , T. (2017). *Veranstaltungen 4.0: Konferenzen, Messen und Events im digitalen Wandel.* Springer Fachmedien Wiesbaden GmbH. doi:DOI 10.1007/978-3-658-16223-8

Kronlachner, M., & Zotter, F. (2013). Spatial transformations for the enhancement of Ambisonic recordings. Von http://www.matthiaskronlachner.com/wp-

content/uploads/2013/01/ICSA2014_KronlachnerZotter_AmbisonicTransformations.pdf abgerufen

Li, Z.-N., Drew, M., & Liu, J. (2014). *Fundamentals of Multimedia: Second Edition.* Springer International Publishing Switzerland. doi:DOI 10.1007/978-3-319-05290-8

Lamapoll.de. (2022). Abgerufen am 24. August 2022 von Online Umfrage erstellen: https://www.lamapoll.de

Mack , S., & Rayburn, D. (2006). *Hands-On Guide to Webcasting: Internet Event and AV Production.* Elsevier Inc. .

Magnor, M., & Sorkine-Hornung , A. (2020). *Real VR – Immersive Digital Reality: How to Import the Real World into Head-Mounted Immersive Displays.* Springer Nature Switzerland AG. doi:https://doi.org/10.1007/978-3-030-41816-8

Malham, D. (1990). Ambisonics-A Technique for Low-Cost, High-Precision, Three-Dimensional Sound Diffusion. (M. T. Group, Hrsg.) S. 118-120. Von https://quod.lib.umich.edu/i/icmc/bbp2372.1990.030/1 abgerufen

Meinel, C., & Sack, H. (2012). *Internetworking: Technische Grundlagen und Anwendungen.* Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-540-92940-6

Meinel, C., & Sack, H. (2014). *Digital Communication: Communication, Multimedia, Security.* Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-642-54331-9

Meta. (2021). Abgerufen am 24. August 2022 von Introducing Meta: A Social Technology Company: https://about.fb.com/news/2021/10/facebook-company-is-now-meta/

Meta. (2021). Abgerufen am 24. August 2022 von META QUEST 2: https://store.facebook.com/de/quest/products/quest-2/tech-specs#tech-specs

Meta. (2022). Abgerufen am 24. August 2022 von Chat-Moderationseinstellungen: https://www.facebook.com/business/help/631291454317308?id=648321075955172

Moreau, S., Daniel, J., & Bertet, S. (20–23. Mai 2006). 3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of a 4th Order Spherical Microphone. *Audio Engineering Society*, S. 24.

Newtek. (2022). Abgerufen am 24. August 2022 von NDI (Network Device Interface) Overview: https://support.newtek.com/hc/en-us/articles/217662358-NDI-Network-Device-Interface-Overview

Obsproject. (2022). Abgerufen am 24. August 2022 von Surround Sound Guide: https://obsproject.com/cs/kb/surround-sound-guide

Paterson, J., & Lee, H. (2021). *3D Audio.* Taylor and Francis. Von https://www.perlego.com/book/2527985/3d-audio-pdf>. abgerufen

Pfanzagl-Cardone, E. (2020). *The Art and Science of Surround and Stereo Recording Including 3D Audio.* Salzburg: Springer-Verlag.

Plugins.iem. (2021). Abgerufen am 24. August 2022 von IEM Plug-in Suite: https://plugins.iem.at

Pro.sony. (2022). Abgerufen am 24. August 2022 von PXW-X70: https://pro.sony/de_DE/products/handheld-camcorders/pxw-x70

Reaper. (2022). Abgerufen am 24. August 2022 von REAPER | Audio Production Without Limits: https://www.reaper.fm

Richards, P. (kein Datum). *The Unofficial Guide to Open Broadcaster Software: OBS: The World's Most Popular Live Streaming Software.* Von https://cpfumc.org/wp-content/uploads/2022/03/The-Unofficial-Guide-to-Open-Broadcaster-Software-PDF.pdf abgerufen

Rode. (2022). Abgerufen am 24. August 2022 von NT-SF1 Ambisonic Microphone: https://rode.com/de/microphones/360-ambisonic/nt-sf1

Roginska , A., & Geluso, P. (2018). *The Art and Science of Binaural and Multi-Channel Audio.* New York: Taylor & Francis.

Schmidt, U. (2013). *Professionelle Videotechnik: Grundlagen, Filmtechnik, Fernsehtechnik, Geräte- und Studiotechnik in SD, HD, DI, 3D.* Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-642-38992-4

Schoeffler, M., Silzle, A., & Herre, J. (8. August 2015). Evaluation of Spatial/3D Audio: Basic Audio Quality vs. Quality of Experience. *JOURNAL OF LATEX CLASS FILES* , S. 14. doi:DOI 10.1109/JSTSP.2016.2639325

Sennheiser. (2022). Abgerufen am 24. August 2022 von AMBEO VR Mic: https://de-de.sennheiser.com/mikrofon-3d-audio-ambeo-vr-mic

Stage entertainment. (2022). Abgerufen am 24. August 2022 von Blue Man Group: https://www.stage-entertainment.de/musicals-shows/blue-man-group-berlin

Stage entertainment. (2022). Abgerufen am 24. August 2022 von Stage Bluemax Theater: https://www.stage-entertainment.de/theatervermietung/stage-bluemax-theater-berlin

STEAM. (2022). Abgerufen am 24. August 2022 von DeoVR Video Player: https://store.steampowered.com/app/837380/DeoVR_Video_Player/

Thakur, A. (2017). *Avidblogs*. Abgerufen am 2022. August 2022 von Facebook 360 Spatial Workstation and Pro Tools | HD 12.8.2: https://www.avidblogs.com/facebook-360-spatial-workstation-and-pro-tools-hd-12-8-2/

THALA. (2022). Abgerufen am 24. August 2022 von THALA music: https://www.facebook.com/whatisthala/about

Tilgner, R. (2021). *Elysia.* Abgerufen am 24. August 2022 von Mastering for Spotify, YouTube, Tidal, Amazon Music, Apple Music and other Streaming Services: https://www.elysia.com/mastering-for-spotify-youtube-tidal-amazon-music-apple-music-and-other-streaming-services/

vuze.camera. (2019). *Youtube.* Abgerufen am 24. August 2022 von VR Studio user interface tutorial: https://www.youtube.com/watch?v=Rad5WDsJLIs&t=234s

Vuze.camera. (2020). Abgerufen am 24. August 2022 von Vuze VR Studio: https://vuze.camera/support/vr-software/vuze-vr-studio

Warnecke, H., & Bullinger, H.-J. (1993). *Virtual Reality Anwendungen und Trends.* Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-642-88650-8

Waves. (2017). Abgerufen am 24. August 2022 von Ambisonics Explained: A Guide for Sound Engineers: https://www.waves.com/ambisonics-explained-guide-for-sound-engineers

Weinzierl, S. (2008). *Handbuch der Audiotechnik.* Berlin: Springer-Verlag Berlin Heidelberg. doi:DOI 10.1007/978-3-540-34301-1

Weiss, J., Nolan, J., Hunsinger, J., & Trifonas, P. (2006). *The International Handbook of Virtual Learning Environments.* Netherlands: Springer.

Wikipedia. (2012). Abgerufen am 24. August 2022 von RGB to Y'CrCb conversation: https://en.wikipedia.org/wiki/File:CCD.png#file

Wikipedia. (2021). Abgerufen am 24. August 2022 von Windows Mixed Reality: https://de.wikipedia.org/wiki/Windows_Mixed_Reality

Wikipedia. (2022). Abgerufen am 24. August 2022 von SteamVR: https://de.wikipedia.org/wiki/SteamVR

Wikipedia. (2022). Abgerufen am 24. August 2022 von Open Broadcaster Software (OBS): https://de.wikipedia.org/wiki/Open_Broadcaster_Software

Windscheid, J., & Gold, B. (2022). *360°-Videos in der empirischen Sozialforschung: Ein interdisziplinärer Überblick zum Einsatz von 360°-Videos in Forschung und Lehre.* Springer Fachmedien Wiesbaden GmbH. doi:https://doi.org/10.1007/978-3-658-34364-4

YouTube. (2022). Abgerufen am 24. August 2022 von 2D Videoschnitt und 360° VR Videoschnitt der Blue Man Group Show 2022: https://www.youtube.com/playlist?list=PLAJeV5TUn-_03_Js_Xpg39PQ0tnceBS9E

YouTube VR. (2022). Abgerufen am 24. 08 2022 von https://vr.youtube.com/

Youtube-Hilfe. (2022). Abgerufen am 24. August 2022 von Grundlagen des Livechats auf YouTube: https://support.google.com/youtube/answer/2524549?hl=de

Youtube-Hilfe. (2022). Abgerufen am 24. August 2022 von Von YouTube empfohlene Einstellungen für die Upload-Codierung: https://support.google.com/youtube/answer/1722171?hl=de#zippy=%2Cconta iner-mp%2Caudio-codec-aac-lc%2Cvideo-codec-h%2Cframerate%2Cvideoaufl%C3%B6sung-und-seitenverh%C3%A4ltnis%2Cbitrate

Youtube-Hilfe. (2022). Abgerufen am 24. August 2022 von Live streaming latency: https://support.google.com/youtube/answer/7444635?hl=en

YouTube-Hilfe. (2022). Abgerufen am 24. August 2022 von Videoauflösung und Seitenverhältnisse: https://support.google.com/youtube/answer/6375112?hl=de&ref_topic=9257782

YouTube-Hilfe. (2022). Abgerufen am 24. August 2022 von Raumklang in 360°-Videos und VR-Videos verwenden: https://support.google.com/youtube/answer/6395969?hl=de#

ZOOM CORPORATION. (2022). *zoomcorp.* Von https://zoomcorp.com/media/documents/D_F6.pdf abgerufen

Zoomcorp. (2022). Abgerufen am 24. August 2022 von DER ZOOM H3-VR: https://zoomcorp.com/de/de/handy-recorder/handheld-recorders/h3-vr-360-audio-recorder/

Zotter, F., & Frank, M. (2019). *Ambisonics A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality.* Cham, Switzerland: Springer Nature Switzerland AG. doi:https://doi.org/10.1007/978-3-030-17207-7

Anhang

A. Einleintung zum Fragebogen

1. Einleitung

Textblock

Liebe Teilnehmende,

durch die Auswirkungen der Corona-Pandemie kam es zu weitreichenden Schließungen von Theatern, Clubs und Konzerthallen. Hierdurch wurden viele Veranstaltungen verschoben oder abgesagt. Als Alternative fanden Digitale-Veranstaltungen im Streaming-Format auf zahlreichen Online-Plattformen statt. Diese wurden größtenteils in herkömmlichen 2D-Videoformaten umgesetzt. Auch die Einbindung von Virtual Reality hat hierdurch deutlich zugenommen und an Bedeutung gewonnen. Daher soll in dieser Studie untersucht werden, wie immersive Videoaufnahmen im direkten Vergleich zu herkömmlichen 2D Videoformaten auf die Zuschauer:innen wirken und wie sie diese bewerten.

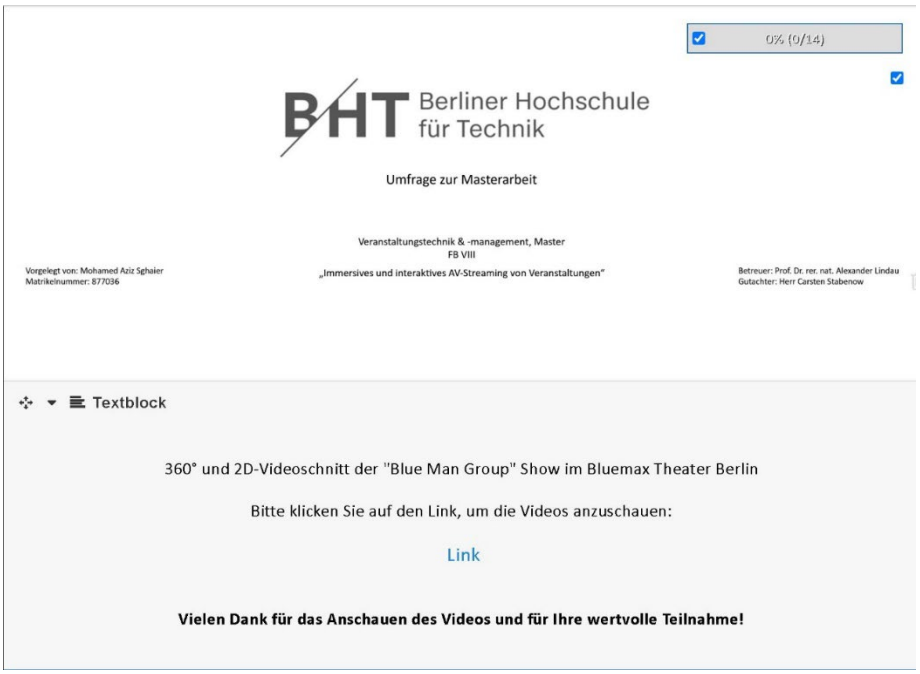
Mit der Hilfe des vorliegenden Fragebogens soll ein ausführliches Meinungsbild der unterschiedlichen Teilnehmenden erstellt werden. Hierbei wird untersucht, wie sich die Nutzung von VR-Inhalten auf das Publikum auswirkt, ob flächendeckende VR-Aufnahmen von Veranstaltungen vermehrt eingesetzt werden sollten und diese eine zukünftige Alternative zu konventionellen 2D Videoformaten darstellen könnten. Für die Durchführung der Umfrage benötigen Sie einen Kopfhörer, um das binaurale 3D-Klangerlebnis optimal wahrnehmen zu können. Die Verwendung einer VR-Brille ist ebenfalls ratsam, um ein vollständig immersives 360°-Erlebnis zu genießen.
Wenn Sie keine VR-Brille besitzen, können Sie einfach das 360°-Video mit der Maus Ihres Computers rotieren lassen, um die aufgenommene Show aus verschiedenen Blickwinkeln zu betrachten.

Sie benötigen ca. 20 Minuten, um den Fragebogen auszufüllen. Der Fragebogen ist anonym und wird nur für das Masterarbeitsprojekt verwendet.

Vielen Dank für Ihre Unterstützung und die Beantwortung der untenstehenden Fragen.

Berliner Hochschule für Technik (BHT)
Mohamed Aziz Sghaier (877036)
Einbecker Str. 23, 10317 Berlin
s74864@beuth-hochschule.de
+4917627027171

B. Fragebogen Seite 2



0% (0/14)

Berliner Hochschule für Technik

Umfrage zur Masterarbeit

Veranstaltungstechnik & -management, Master
FB VIII
„Immersives und interaktives AV-Streaming von Veranstaltungen"

Vorgelegt von: Mohamed Aziz Sghaier
Matrikelnummer: 877036

Betreuer: Prof. Dr. rer. nat. Alexander Lindau
Gutachter: Herr Carsten Stabenow

Textblock

360° und 2D-Videoschnitt der "Blue Man Group" Show im Bluemax Theater Berlin

Bitte klicken Sie auf den Link, um die Videos anzuschauen:

Link

Vielen Dank für das Anschauen des Videos und für Ihre wertvolle Teilnahme!

C. Fragebogen Seite 3

25% (4/14)

Berliner Hochschule für Technik

Umfrage zur Masterarbeit

Veranstaltungstechnik & -management, Master
FB VIII
„Immersives und interaktives AV-Streaming von Veranstaltungen"

Vorgelegt von: Mohamed Aziz Sghaier
Matrikelnummer: 877036

Betreuer: Prof. Dr. rer. nat. Alexander Lindau
Gutachter: Herr Carsten Stabenow

5. Wie fanden Sie die Bildqualität der 360°-Videoaufnahme im Vergleich zu 2D-Videoaufnahme?

Viel schlechter | Etwas schlechter | Gleich | Etwas besser | Viel besser

6. Wie fanden Sie die 3D-Soundqualität der 360°-Videoaufnahme im Vergleich zu 2D-Videoaufnahme?

Viel schlechter | Etwas schlechter | Neutral | Etwas besser | Viel besser

7. Wie sehr haben Sie sich beim Betrachten des 360°-Videos im Vergleich zu 2D-Aufnahme visuell umhüllt gefühlt?

Gar nicht | Umhüllungsgefühl nicht besonders bemerkbar | Neutral

Etwas umhüllt | Komplett umhüllt

D. Fragebogen Seite 4

8. Wie sehr haben Sie sich beim Anhören des Sounds des 360°-Videos im Vergleich zur 2D-Aufnahme auditiv umhüllt gefühlt?

- Gar nicht
- Umhüllungsgefühl nicht besonders bemerkbar
- Neutral
- Etwas umhüllt
- Komplett umhüllt

9. Wie stark war das Gefühl vor Ort zu sein bei der 360°-Videoaufnahme?

- Gar nicht
- Nicht besonders
- Normal
- Etwas
- Sehr stark

10. Wie stark war das Gefühl vor Ort zu sein bei der 2D-Videoaufnahme?

- Gar nicht
- Nicht besonders
- Normal
- Etwas
- Sehr stark

‹ Zurück Weiter ›

E. Fragebogen Seite 5

11. Zur Zeit kann man ein gutes VR-Headset, wie z.B das Oculus Quest 2, für 349 € bestellen. Würden Sie sich für diesen Preis ein VR-Headset kaufen?

- Ja
- Nein
- Vielleicht

12. Wie mehr (oder auch weniger) wären Sie bereit, für eine Online-Veranstaltung in gezeigten 3D-VR-Variante zu bezahlen (z. B. einen virtuellen Theater- oder Konzertbesuch)?

- -10 € oder weniger
- -5 €
- -1 €
- 0 €
- 1 €
- 5 €
- 10 € oder mehr

13. Haben Sie sich während des Betrachtens des VR-Videos unwohl gefühlt?
Wenn ja, wählen Sie bitte die passende Beschreibung (Mehrfachantworten möglich)

- [ ] Übelkeit
- [ ] Schwindel
- [ ] Unbequem
- [ ] Sehstörungen
- [ ] Hörstörung
- [ ] Etwas anderes. Bitte beschreiben Sie es im untenstehenden Kommentarfeld!

14. Bitte kommentieren Sie hier, was Ihnen im Zusammenhang mit Ihren abgegebenen Bewertungen evtl. aufgefallen ist/was Sie uns näher erläutern bzw. mitteilen wollen:

- Anmerkung

‹ Zurück Beenden ›