

D I P L O M A R B E I T

Nachhallfreie Audioproduktion und
stereoskopische Videoproduktion und
Postproduktion eines Streichquartetts

Durchführung: Fachgebiet Audiokommunikation
Institut für Sprache und Kommunikation

Studiengang: Elektrotechnik

vorgelegt von: Andreas Steckmann

■■■■■■■■■■ ■■■■■■

Gutachter: Prof. Dr. Stefan Weinzierl
Prof. Dr.-Ing. Adam Wolisz

Abgabetermin: 24.07.2014

Eidesstattliche Erklärung

Die selbständige Anfertigung der vorliegenden Arbeit versichere ich an Eides statt.

Berlin, 24. Juli 2014

Andreas Steckmann

Alle erwähnten Firmen- und Markennamen gehören ihren Eigentümern.

Inhaltsverzeichnis

1 Einleitung	1
1.1 Motivation	1
1.2 Stand der Forschung	3
1.3 Die CAVE an der TU Berlin - eine plausible, stereoskopische, virtuelle Umgebung	5
2 Grundlagen des stereoskopischen Sehens	7
2.1 Physiologische Grundlagen	7
2.2 Stereoskopisches Sehen	9
2.3 Wahrnehmung der Raumtiefe.....	11
2.4 Raumwahrnehmung für die stereoskopische Videoproduktion und die CAVE	15
3 Technische Realisation einer virtuellen Umgebung	17
3.1 Grundlagen und Grenzen der visuellen Wahrnehmung	17
3.1.1 Sichtbares Licht	17
3.1.2 Hellempfindlichkeit und Empfindungsbereich.....	18
3.1.3 Farbwahrnehmung.....	20
3.1.4 Sensusmotorik.....	22
3.1.5 Örtliches Auflösungsvermögen	23
3.1.6 Zeitliches Auflösungsvermögen	25
3.2 Adaption der Technik an die visuelle Wahrnehmung	26
3.2.1 Optische Aufnahmesysteme und Parameter.....	26
3.2.2 Optische Wiedergabesysteme und Parameter.....	35
3.2.2.1 Aktive Projektionssysteme	37
3.2.2.2 Passive Projektionssysteme.....	38
3.2.3 Verfahren der Stereoskopie	40
3.2.3.1 Verfahren für stereoskopische Filmaufnahmen	41
3.2.3.2 Verfahren für stereoskopische Filmwiedergabe	42
3.3 Grundlagen und Grenzen der auditiven Wahrnehmung.....	48
3.3.1 Schall.....	48
3.3.2 Physiologische Grundlagen	50
3.3.3 Auditiv Wahrnehmung.....	51
3.3.4 Auditiv Lokalisation	53
3.3.5 Auditiv Wahrnehmung in geschlossenen Räumen	55
3.4 Adaption der Technik an die auditive Wahrnehmung.....	57
3.4.1 Akustisches Aufnahmesystem für eine CAVE.....	57
3.4.1.1 Mikrofoning zur akustischen Akquise	57

3.4.1.2 FABIAN – ein System zur automatisierten Akquise kopfbezogener BRIRs	59
3.4.2 Akustisches Wiedergabesystem für eine CAVE	61
4 Nachhaltfreie Audioproduktion eines Streichquartetts	64
4.1 Vorbereitung	64
4.2 Durchführung.....	66
4.3 Postproduktion.....	71
4.4 Fehler und Hemmnisse bei nachhaltfreien Audioproduktionen.....	73
5 Stereoskopische Videoproduktion und Postproduktion eines Streichquartetts	74
5.1 Ermittlung der Perspektiven.....	75
5.2 Vorbereitung	78
5.3 Durchführung.....	79
5.4 Postproduktion und Visual Composing	84
5.5 Fehlerquellen und Probleme bei stereoskopischen Videoproduktionen	93
6 Integration der Streichquartett-Panoramen in die CAVE.....	95
7 Zusammenfassung und Ausblick.....	98
Abkürzungen	100
Quellenverzeichnis.....	102
Anhang.....	108
Abbildungsverzeichnis	121

1 Einleitung

1.1 Motivation

Die Fähigkeit zur dreidimensionalen Wahrnehmung des Raumes ist ein wichtiges Resultat der evolutionären Entwicklung des Menschen, das in großem Maße zu seinem Überleben beigetragen hat. Mit Hilfe der zwei Augen, deren Sichtfelder sich überlagern ist das Gehirn in der Lage, ein visuelles, räumliches Abbild der Umwelt zu konstruieren und eine genaue Abschätzung der Entfernung zu Objekten zu ermöglichen. Ebenso ist die Ortung und räumliche Wahrnehmung von Schallwellen mit Hilfe der Ohren und die neuronale Weiterverarbeitung der daraus resultierenden auditiven Signale möglich. Obwohl der Mensch fünf Sinne besitzt, werden zur räumlichen Erfassung der Umwelt im Wesentlichen der dominierende visuelle und der auditive Sinn eingesetzt. In unmittelbarer Umgebung spielt der taktile (Tast) Sinn noch eine Rolle, jedoch nicht bei Distanzen, die über die Armlänge, die natürliche Grenze taktiler Erfassungen, hinausgehen.

Um die im Fokus stehende Raumwahrnehmung unter Laborbedingungen untersuchen zu können, bietet sich die Schaffung einer virtuellen Umgebung an. Je höher in dieser so genannten virtuellen Realität (VR¹) der Grad der Immersion –das Gefühl des „Eintauchens“– für die Probanden ist, umso plausibler ist die VR. Da der Mensch seine Informationen für die Wahrnehmung aus den visuellen und der auditiven Sinneseindrücken bezieht, muss eine virtuelle Umgebung so gestaltet sein, dass die erzeugten Informationen für die Sinneseindrücke denen der Realität möglichst gleichen [1, S. 84]. Die Generierung dieser die Realität repräsentierenden Informationen stellt eine große Herausforderung an eine technische Realisierung einer VR dar, insbesondere hinsichtlich der vom visuellen Sinn zu erfassenden 50 Milliarden bzw. der vom auditiven Sinn zu erfassenden eine Million Bit je Sekunde an Information [2, S. 13].

Die DFG-Forschergruppe SEACEN² hat sich im Rahmen des Teil-Projektes P9 “audio-visual perception of acoustical environments”³ das Ziel gesetzt eine möglichst plausible, virtuelle Umgebung zu schaffen, um den Einfluss akustischer und optischer Eigenschaften von Konzertsälen als Stimuli für die visuelle und auditive Raumwahrnehmung zu untersuchen. Dazu stellt die Forschergruppe am Fachgebiet Audiokommunikation der TU Berlin eine optoakustische virtuelle Umgebung - eine CAVE⁴- her. Die CAVE besteht aus einem stereoskopischen 180° umfassenden Panorama-Projektionssystem, das mit Binauraltechnik⁵ zu einer

¹ Virtual Reality

² Simulation and Evaluation of Acoustical Environments – Simulation und Evaluierung von akustischen Umgebungen

³ audio-visuelle Wahrnehmung von akustischen Umgebungen

⁴ CAVE - Cave Automatic Virtual Environment – mittels Computer simulierte künstlich erschaffene Umgebung

⁵ Verfahren zur Reproduktion realer Schallfelder [2, S. 14]

plausiblen virtuellen Umgebung kombiniert wurde (Abbildung 1). Darin werden den Probanden sechs akustische und optische virtuelle Räume dargeboten, die unabhängig voneinander variiert werden können. Durch die Integration der stereoskopischen Halbpanoramen von sechs Konzertsälen mit deren jeweiliger binauraler¹ Klangcharakteristik und einem im Vordergrund befindlichen stereoskopisch aufgenommenen Streichquartett soll eine plausible Immersion bei den Probanden erzeugt werden. Des Weiteren können insbesondere *cross modale*² Effekte zwischen visueller und auditiver Wahrnehmung, die beispielsweise bei der Kombination eines stereoskopischen Konzertsaal-Hintergrundbildes mit der Binaural-Charakteristik eines anderen Konzertsaaes auftreten, untersucht werden (Abbildung 2). Die Auswertung der physikalischen und perzeptiven Messergebnisse soll bei der Schaffung eines empirischen Modells helfen. Mit dessen Hilfe sollen die Parameter der Raumwahrnehmung auf der Grundlage der physikalischen Raumeigenschaften prognostiziert werden können. Die gewonnenen Erkenntnisse sollen zur Erhöhung des Verständnisses von intramodalen Prozessen der Wahrnehmung beitragen. Sie sind eine wichtige Voraussetzung für die Weiterentwicklung der Simulationstechnik und der Gestaltung von audio-visuellen, virtuellen Umgebungen [3, S. 189 -190].

Die vorliegende Arbeit führt zunächst in die Grundlagen des menschlichen stereoskopischen Sehens, in grundlegende stereoskopische Aufnahme- und Wiedergabeverfahren in der Videotechnik und in die technischen Grundlagen zur Realisation einer virtuellen Umgebung ein. In Kapitel 4 wird die Schaffung von technischen Voraussetzungen und die Durchführung der Audioaufnahmen eines Streichquartetts im reflexionsarmen Raum am Institut für Technische Akustik der TU Berlin dokumentiert. Die Aufnahmen dieser nachhallfreien Audioproduktion werden anschließend mit den bereits akquirierten Impulsantworten der sechs Konzertsäle gefaltet und können so den Ohren als binaurales, plausibles Schallfeld dargeboten werden. In Kapitel 5 wird die stereoskopische Greenscreen-Videoaufnahme³ des

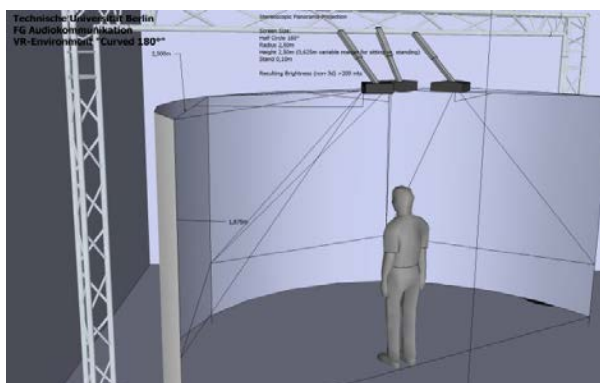


Abbildung 1: Prinzip der CAVE an der TU Berlin
[3, Abb. S. 209]

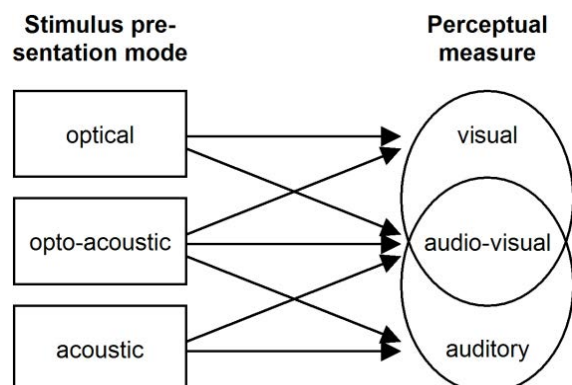


Abbildung 2: Untersuchung cross modaler Effekte
[3, Abb. S. 211]

¹ lat. „mit beiden Ohren“

² Überschneidungseffekte

³ Film-/ Videoaufnahmetechnik zur Erzeugung eines transparenten Bildhintergrundes

Streichquartetts erläutert. Unter Beachtung von Perspektiventoleranz und korrekter Montage von stereoskopischem Streichquartett-Greenscreen-Video mit den Konzertsaal-Panoramen folgt die Darstellung der stereoskopischen Video-Postproduktion. Durch synchrone, aber variable Kombination der binauralen Audioaufnahmen der sechs Konzertsäle mit dem im Vordergrund befindlichen Streichquartett-Video und den stereoskopischen Hintergrundpanoramen der sechs Konzertsäle entstand die in Kapitel 6 beschriebene Integration zu einer virtuellen Umgebung - der CAVE. Nach einer Zusammenfassung wird ein kurzer Ausblick in die weitere Entwicklung und mögliche Anwendungsbereiche des vorliegenden variablen, binauralen, stereoskopischen 180° umfassenden Panorama-Projektionssystems gegeben.

1.2 Stand der Forschung

Um eine virtuelle Realität mit einer plausiblen Immersion zu schaffen, ist es notwendig, den Wahrnehmungsprozess zu verstehen. Die in diesem Zusammenhang auftretenden Fragen grenzen z.T. an Bereiche der Philosophie in Bezug auf grundsätzliche Fragestellungen wie: Was ist Realität? Was nimmt der Mensch mittels seiner Sinne von der Umwelt wahr? Wie kann die erfahrbare Realität mittels Technik so realistisch wie möglich abgebildet werden? Um die komplexen Mechanismen der menschlichen Wahrnehmung zu verstehen, wurde ein Modell entwickelt, das den Wahrnehmungsprozess in einzelne Schritte untergliedert (Abbildung 3). Die kreisförmige Anordnung veranschaulicht die Dynamik und die ständigen Veränderungen des Prozesses [4, S. 4]. Die vom Menschen erfassten Reize (beachtete Stimuli) erregen die Rezeptoren der Sinnesorgane. Auf der Netzhaut (Retina) des Auges entsteht ein Abbild des beachteten optischen Reizes, und der von den Gehörknöchelchen auf das Trommelfell auftreffende Schall wird als erregende Schwingung weiter ins Innenohr geleitet. Die Sinneszellen wandeln die Erregungen in elektrische Signale um (Transduktion), die über neuronale Netze an das Gehirn weitergeleitet werden. Im Seh- bzw. Hörzentrum erfolgt die neuronale Verarbeitung der in elektrische Signale transformierten Stimuli (Abbildung 4). Die Auswertung der zahlreichen neuronalen Signale ermöglicht schließlich eine bewusste, sensorische Erfahrung – die Wahrnehmung. Um die wahrgenommenen Reize in zutreffende Kategorien einordnen zu können, ist *Erkennen* notwendig. Der Unterschied zwischen Wahrnehmung und Erkennen wird bei der Betrachtung von Dr. P., dessen Erkennen nur fehlerhaft funktionierte, deutlich. So wurde von ihm eine wahrgenommene Parkuhr fälschlicherweise als Mensch erkannt, obwohl bei ihm keine Beeinträchtigung im Sehen bestand. Die Ursache war ein Hirntumor, der die Fähigkeit, bekannte Objekte korrekt zu erkennen bzw. einzuordnen, verhinderte (Agnosie). Während der neuronalen Verarbeitung werden die aufgenommenen Informationen mit bereits vorhandenen im Gehirn

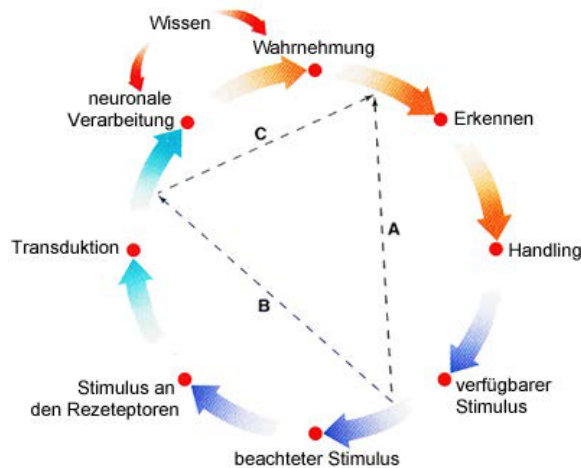


Abbildung 3: Wahrnehmungsprozess
[4, Abb. S. 4]

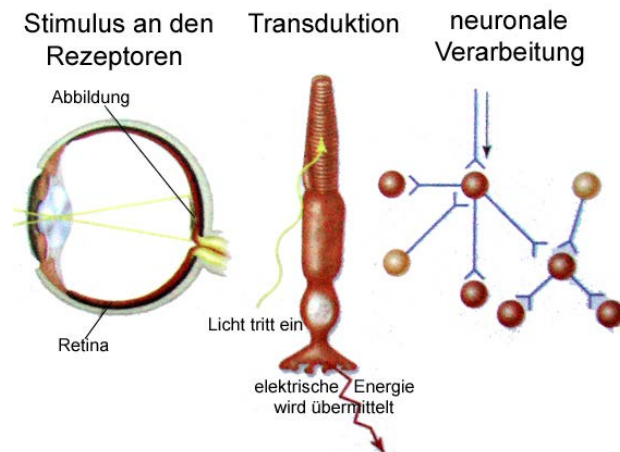


Abbildung 4: drei Teilschritte des Wahrnehmungsprozesses [4, Vgl. Abb. S. 5]

gespeicherten Informationen bzw. vorhandenem Wissen verglichen und können so genauer klassifiziert und bewertet werden. Als wichtiges Resultat des Wahrnehmungsprozesses wird von einigen Forschern *Handlung* insbesondere zur Sicherung des Überlebens angesehen. Ein genaues Abbild der Umwelt schafft Voraussetzungen zur Handlung und ermöglicht dadurch dem Menschen, wie auch anderen Lebewesen, eine optimale Anpassung an die Umwelt. Da sich die Umwelt bzw. der Bezug zu ihr (z.B. durch eine vorangegangene Handlung) ständig ändert, ist eine permanente Wahrnehmung der Umwelt und eine entsprechend angepasste Handlung notwendig. Daraus resultiert die kreisförmige Anordnung aller Teilschritte des Wahrnehmungsprozesses (Abbildung 3) [4, S. 4-7].

Diese Erkenntnisse aus der Wahrnehmungspsychologie wurden bzw. werden z.T. bei der Erzeugung virtueller Umgebungen berücksichtigt. So ist neben der Erzeugung von stereoskopischen Bildern und Filmen mit immer besserer Qualität und Auflösung, die Synchronität sowie eine hohe zeitliche Auflösung von Filmen wichtig. Insbesondere sollte die virtuelle Umgebung *Handlung* als Reaktion auf die (virtuelle) Wahrnehmung erlauben, ohne dass dabei Beeinträchtigungen der Immersion entstehen. Das war bei den seit Anfang der 70er bis Mitte der 80er Jahre des 20. Jh. entwickelten HMDs¹ nicht immer der Fall. Dabei sind zwei direkt vor den Augen des Benutzers positionierte kleine Anzeigen in einem helmähnlichen Gehäuse untergebracht [1, S. 87]. Während sich der Benutzer durch Kopfbewegungen in der virtuellen Welt umsehen kann (Handlung), werden aufgrund des HMD-Gewichtes der taktile und auch der Gleichgewichtssinn gereizt, wodurch die Immersion gestört werden kann. Die durch diese und weitere Nachteile wie eine geringe Bildauflösung und schnelles Ermüden gekennzeichneten HMD-Systeme wurden zunehmend durch Projektionssysteme mit größeren Darstellungen in hoher Auflösung verdrängt [1, S. 87].

¹Head Mounted Display – eine mit dem Kopf verbundene Anzeige

Anfang der 90er Jahre des 20. Jh. wurde in Chicago an der University of Illinois die CAVE entwickelt (Abbildung 5). Diese mittels 3D Projektoren und leistungsfähigen, bildverarbeitenden Computern auf Leinwänden realisierte würfelförmige virtuelle Umgebung befreite die Anwender von den schweren HMD-Systemen und machte eine neue Art der virtuellen Realität erfahrbar (Abbildung 6). Obwohl VR-Anwendungen in Bereichen wie Automobilindustrie, Medizin oder Architektur eingesetzt werden [5, S. 9-12], sind Zusammenhänge bzw. Wechselwirkungen zwischen Sehen und Hören bislang nur wenig erforscht. Es existieren zahlreiche Studien im Bereich Wahrnehmung von Raumakustik, die sich jedoch darauf konzentrieren, technisch messbare („objektive“) Raumeigenschaften zu finden mit deren Hilfe Voraussagen über („subjektive“) Raumwahrnehmungen wie Raumeindruck, auditiv wahrgenommene Raumgröße oder Präsenz möglich sind. Nur wenige Studien beziehen sich auf die technische Simulation von Realität in der Raumakustik. Die bisher vorwiegend empirisch untersuchten audio-visuellen Wechselwirkungen sind im Wesentlichen noch Teil der Grundlagenforschung [3, S. 189-192]. Eine der wenigen vergleichbaren CAVE-Anwendungen wurde an der RWTH Aachen umgesetzt [8]. Mit drei Teil-Projekten zu raumakustischen Themen u.a. „Perceptual-based optimization of room acoustic auralization parameters“ ist die RWTH Aachen ebenfalls in das SEACEN-Projekt mit eingebunden [9].

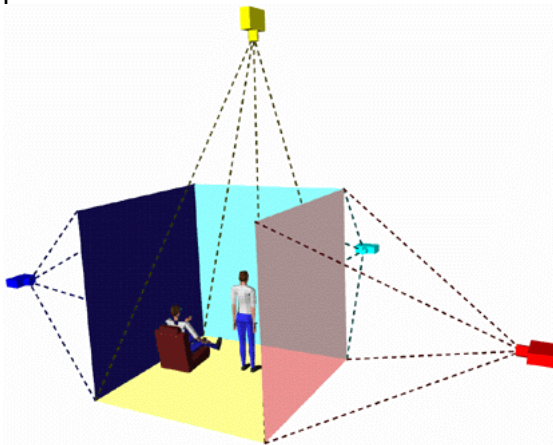


Abbildung 5: Prinzip-Aufbau einer CAVE [6]

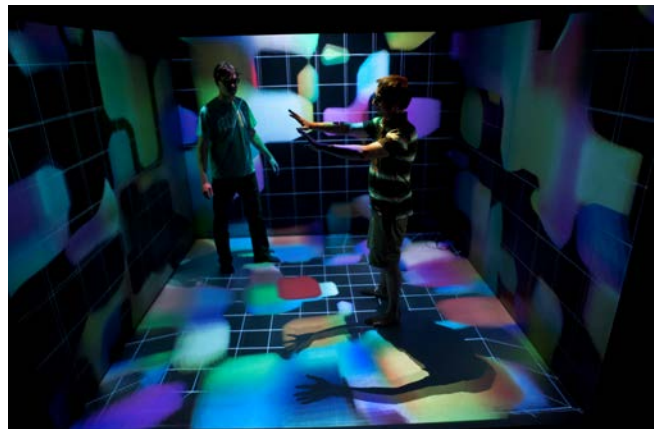


Abbildung 6: in einer CAVE [7]

1.3 Die CAVE an der TU Berlin - eine plausible, stereoskopische, virtuelle Umgebung

Das Ziel von VR Anwendungen besteht darin, Sinneswahrnehmungen zu erzeugen, die eine plausible Immersion, also ein vollständiges Eintauchen in die virtuelle Welt, ermöglichen [1, S. 84]. Dabei liegt die auf Teilaspekte fokussierte Forschungsstrategie beim SEACEN-Projekt, im Gegensatz zu den wenigen zum Thema *Audio Visuelle Raumwahrnehmung* existierenden Studien, auf einem integrierten, opto - akustisch ausgewogenen Ansatz. Durch die Schaffung methodischer Voraussetzungen soll die experimentelle Trennung von akustischen und optischen Einflüssen auf die Raumwahrnehmung ermöglicht werden. Dazu werden die, wie in der vorliegenden Arbeit später dargestellt, zuvor akquirierten Daten bzw.

Aufnahmen der musikalischen Aufführung eines Streichquartetts als opto-akustische Reize verwendet. Diese Film- und Audiodaten können mittels Greenscreen-Technik mit den bereits akquirierten stereoskopischen Panorama-Hintergründen von sechs Konzertsaal-Abbildungen zu einer stereoskopischen, virtuellen Umgebung kombiniert werden. Zunächst kann mittels Bewertung durch Probanden die Plausibilität der CAVE eingeschätzt werden. Darüber hinaus können durch die unabhängige Variation von akustischen und optischen Räumen, in diesem Fall Konzertsälen, cross-modale Wechselwirkungen erforscht und abgeschätzt werden [3, S. 189 -190].

Um die hohen Anforderungen einer plausiblen Immersion zu erreichen, wurden für die TU-CAVE (Abbildung 7) neben qualitativ hochwertigem Equipment, wie fünf hoch auflösenden, mit hoher Bildfrequenz betriebenen Projektoren und stereoskopischen Shutter-Brillen, insbesondere neue Verfahren für die Daten-Akquise verwendet. Auf diese Verfahren, welche Daten liefern, die eine plausible, virtuelle Umgebung erzeugen können, wird in dieser Arbeit in den Kapiteln 4 zur Audioproduktion und Kapitel 5 zur Videoproduktion eingegangen.

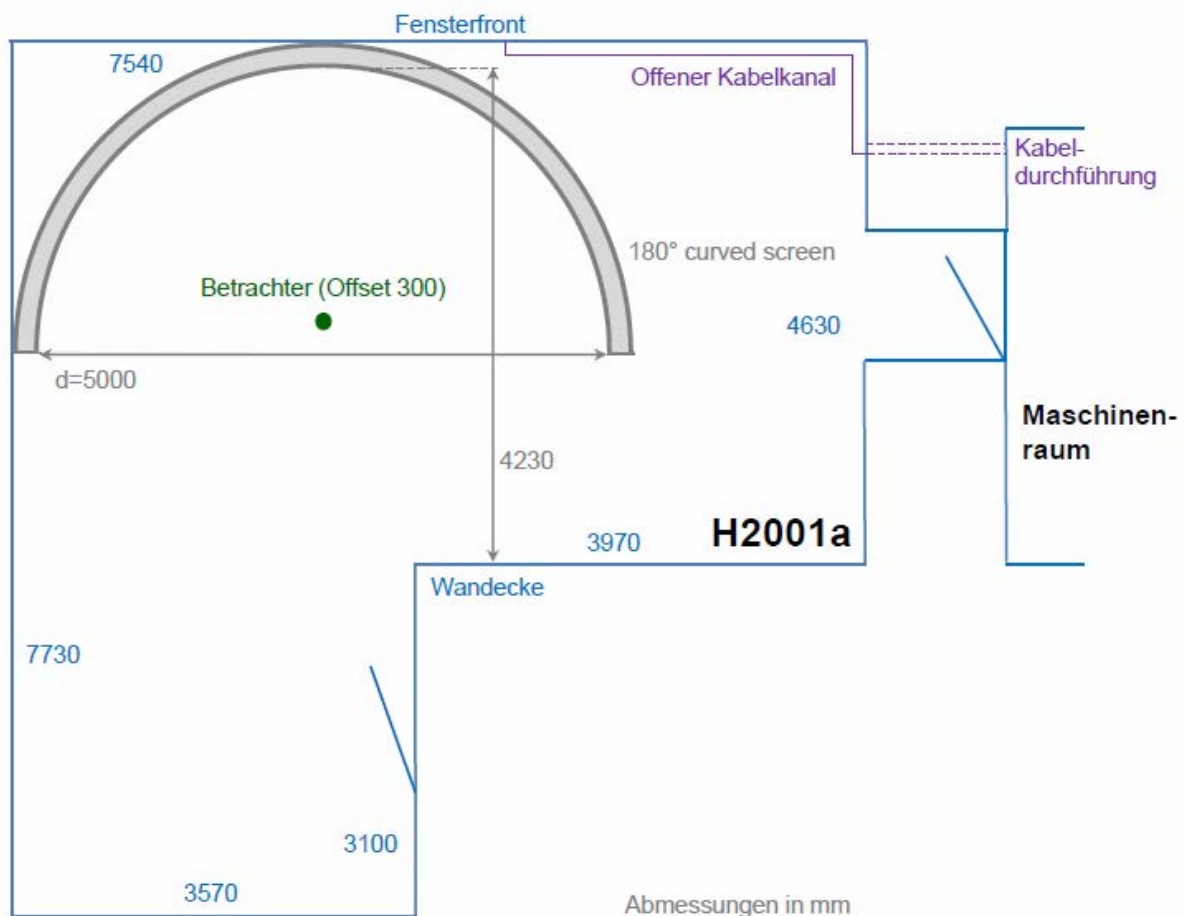


Abbildung 7: Grundriss der CAVE am Fachbereich Audiokommunikation der TU Berlin

2 Grundlagen des stereoskopischen Sehens

Um eine plausible, virtuelle, stereoskopische Umgebung zu erzeugen, sind Kenntnisse über das stereoskopische menschliche Sehen notwendig. Diese Grundlagen sollen im Folgenden dargelegt werden. Einführend sollen dabei die Grundlagen des menschlichen visuellen Systems erläutert werden.

2.1 Physiologische Grundlagen

Das menschliche Auge ist das Sinnesorgan des Menschen, das optische Reize durch Lichteinfall erfasst und zum Gehirn weiterleitet (Abbildung 8). Es ist kugelförmig und hat einen Durchmesser von ca. 2,5 cm. Vom umgebenden Gewebe wird das Auge durch die schützende Lederhaut abgegrenzt. Die darunter befindliche Aderhaut versorgt das Auge mit Nährstoffen. Die unterste Schicht –die Netzhaut– dient dank ihrer lichtempfindlichen Rezeptoren als „Leinwand“ auf der das Abbild der Umwelt entsteht. Der gallertartige Glaskörper hält das Auge in Form und schützt alle empfindlichen Teile. Die außen liegende Hornhaut bricht das einfallende Licht, das nur durch den durchsichtigen Teil der Hornhaut einfallen kann. Durch das Zusammenwirken von Linse und Hornhaut entsteht auf der Netzhaut ein scharfes, verkleinertes, umgekehrtes, spiegelverkehrtes Bild (Abbildung 9). Da sich die Netzhaut nur langsam an Änderungen der Leuchtdichte anpassen kann, regelt die sich durch An- bzw. Entspannung der Muskulatur schnell schließende Regenbogenhaut (Iris) die Menge des einfallenden Lichtes und schützt so die Netzhaut vor zu großer Lichteinstrahlung. Die Iris funktioniert also ähnlich einer Kamerablende auch in der Hinsicht, dass beim Nahsehen eine Verkleinerung der Pupille (Blende) eine Erhöhung der Tiefenschärfe zur Folge hat [10, S. 12-13].

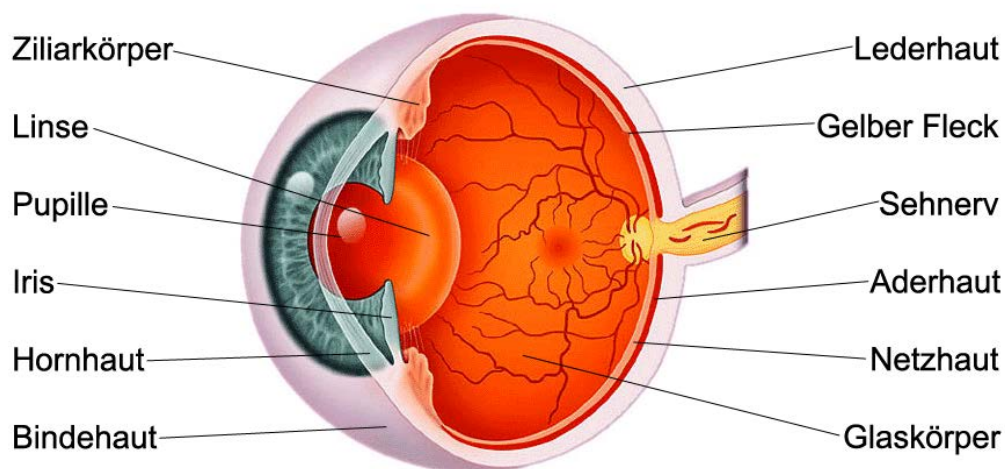


Abbildung 8: Anatomie des menschlichen Auges [11, Vgl. Abb. S. 2]

Die in der Netzhaut vorhandenen Fotorezeptorzellen sind aus lichtempfindlichen Molekülen aufgebaut, die sich bei Lichteinfall strukturell verändern. Entsprechend ihren Eigenschaften sind die Rezeptoren in Zapfen und Stäbchen unterteilt (Abbildung 9, rechts). Während die Stäbchen eine eher geringe Auflösung und weniger scharfes schwarz-weiß Sehen ermöglichen, gewährleisten die Zapfen das Farbsehen, eine hohe Auflösung und Schärfe. Insbesondere sind in der Fovea centralis -dem Punkt des schärfsten Sehens- (kleines Areal innerhalb des Gelben Fleckes) viele Zapfen konzentriert, woraus die Eigenschaft dieser Region –scharfes und hoch auflösendes Farb-Sehen- resultiert. Bei Fixation eines Objektes werden die Augen so gedreht, dass das Objekt auf die Fovea centralis fällt und damit scharf und hoch aufgelöst wahrgenommen werden kann (Abbildung 9, links). In der Peripherie des Sehfeldes auf der Netzhaut werden dagegen hauptsächlich globale und Bewegungsinformationen erfasst [12, S. 50-51]. Die in den Rezeptoren enthaltenen, lichtempfindlichen sogenannten Sehpigmente lösen nach einer Stimulation mit Licht elektrische Signale aus, die durch ein Netzwerk von Neuronen fließen [4, S. 30]. Dabei gelangen die Informationen über eine synaptische Verbindung zu (bipolaren) Nervenzellen. Diese sind wiederum mit Ganglienzellen verbunden, deren Ausläufer (Axone) sich im optischen Nerv bündeln und aus dem Auge heraus weiter zum Gehirn führen. Jede der 800.000 Ganglienzellen encodiert Informationen aus einem kleinen Netzhaut-Areal. Die Feuerrate des Axons einer Ganglienzelle entspricht in ihrem Betrag der auf diese Ganglienzelle auftreffenden Lichtmenge [12, S. 51].

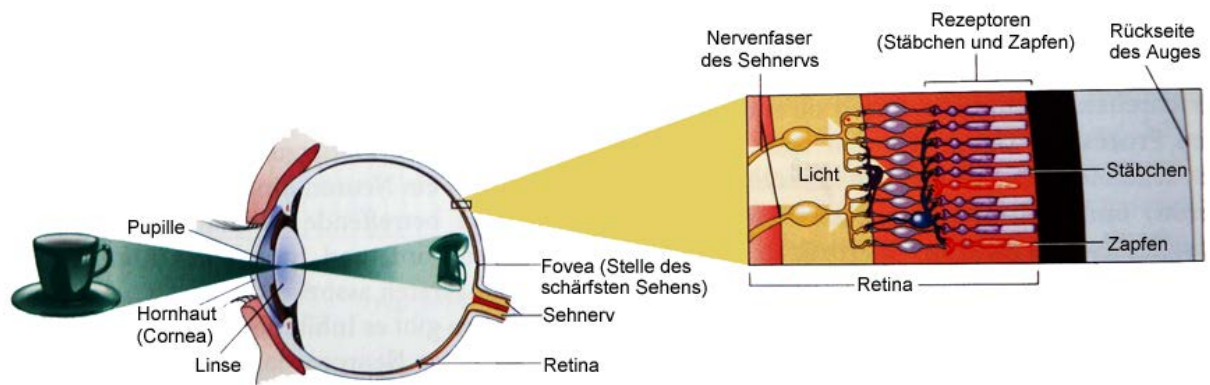


Abbildung 9: Objektabbildung auf der Netzhaut (links); Rezeptoren in der Netzhaut (rechts) [4, Abb. S. 30]

Jede Netzhaut ist senkrecht in eine linke und eine rechte Hälfte geteilt ist. Dementsprechend werden auch die in elektrischen Signalen verschlüsselten Informationen auf geteilten Wegen weitergeleitet. An der Kreuzung der Sehbahnen (Chiasma opticum) verzweigen sich die optischen Nerven beider Augen derart, dass die rechte Hälfte jeder Netzhaut in der rechten Großhirnhemisphäre und die linke Hälfte jeder Netzhaut in der linken Großhirnhemisphäre repräsentiert werden (Abbildung 10). Das ermöglicht dem visuellen System den örtlich direkten Vergleich beider Netzhaut-Teilbilder, was letztlich die Grundlage der Wahrnehmung räumlicher Tiefe darstellt. Der größte Teil der Sehnervenfasern (90%) enden im sogenannten seitlichen Kniehöcker (CGL - Corpus Geniculatum Laterale) im Thalamus des Zwischenhirns

– der ersten höheren Verarbeitungsstufe. In dieser Schaltstation werden Informationen vorgefiltert und z.T. bewertet [10, S. 26]. Vom CGL werden die Informationen über die Sehbahnen weiter zu der im hinteren Schädelbereich befindlichen primären Sehrinde geleitet. In dieser scheckkartengroßen 3 mm dicken primären Sehrinde erfolgt die Datenanalyse. Entsprechende Forschungen an Katzen zeigten, dass sich die Sehrinde in abwechselnde für das linke und das rechte Auge senkrecht durch parallele Schichten verlaufende Blöcke gliedert. In dieser Zauberwürfel – ähnlichen Struktur verarbeitet jeder „Mini-Würfel“ die Signale eines spezifischen Netzhaut-Bereiches (Abbildung 11). Somit wird die gesamte Netzhaut als zusammenhängendes Areal in der Sehrinde repräsentiert. Letztlich entsteht im Gehirn mittels komplex miteinander verbundener Nervenzellen das dreidimensionale Abbild unserer komplexen Umwelt [10, S. 30-31].

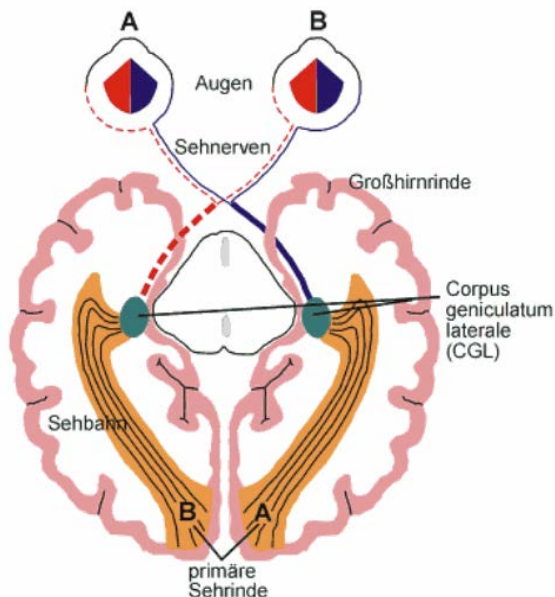


Abbildung 10: Schnitt durch das Gehirn mit den wichtigsten Stationen der Sehbahnen [10, Abb. S. 30]

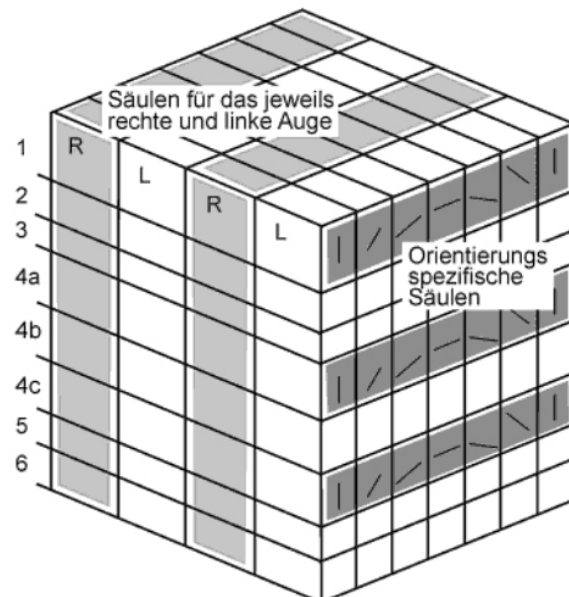
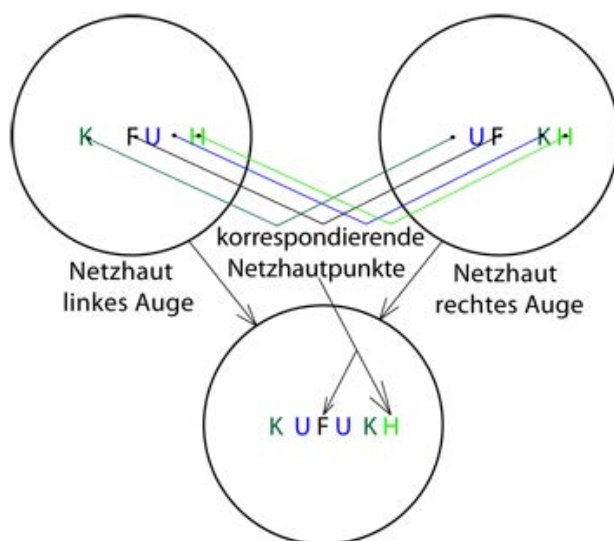


Abbildung 11: Aufbau der Sehrinde [10, Abb. S. 31]

2.2 Stereoskopisches Sehen

Das räumliche Sehen ermöglicht dem Menschen sich optimal in seiner Umwelt zu orientieren und zu bewegen. Insbesondere können Abstände zwischen bzw. Entfernungen von Objekten genau bestimmt werden. Zu den wichtigsten Voraussetzungen des stereoskopischen Sehens gehört die dicht nebeneinander liegende Anordnung der Augen im Abstand von durchschnittlich 6,3 cm, woraus sich ein großer Überlappungsbereich der beiden Gesichtsfelder ergibt. Die daraus resultierende, wichtigste Komponente des visuellen Systems für die Tiefenwahrnehmung ist die *Querdisparation*. Diese führt zu einer durch den Augenabstand bedingten leicht unterschiedlichen Perspektive jedes Auges [1, S. 85]. Damit verbunden ist

eine leicht verschobene Abbildung auf jeder der beiden Netzhäute. Da sich die Sehbahnen so kreuzen, dass die rechte Hälfte jeder Netzhaut in der rechten und die linke Hälfte jeder Netzhaut in der linken Großhirnhemisphäre repräsentiert werden (Abbildung 10; Vgl. Abschnitt 2.1, S. 8), ist eine Überlappung und ein Vergleich der Bilder des rechten und linken Auges und damit die Verarbeitung der stereoskopischen Information am gleichen Ort in der Sehrinde möglich. Bei einem gedachten Übereinanderlegen beider Netzhäute ergibt sich in der Mitte bzw. auf dem Horopter eine Übereinstimmung der Bildinformation währenddessen an den Rändern jeweils unterschiedlich abgebildet wird (Abbildung 12). Dabei sind Punkte, die an den gleichen Stellen jeder Netzhaut liegen als *korrespondierende Netzhautpunkte* mit den jeweils gleichen Stellen im visuellen Cortex¹ verbunden. Die Bildinformationen eines Objektes, die auf beiden korrespondierenden Netzhautpunkten gleich abgebildet werden, liegen auf einer aus dem Strahlengang resultierenden Kreislinie – dem *Horopter*, auf der auch der Fixationspunkt liegt (Abbildung 13). Da diese korrespondierenden Netzhautpunkte jedoch die gleiche Bildinformation besitzen, tragen sie nicht zum stereoskopischen Sehen bei. Im Gegensatz dazu gewinnt das visuelle System die stereoskopische Information aus



Überlagerung von linker und rechter Netzhaut

F – Fixationspunkt auf Horopter

H – nicht fixierter Punkt auf Horopter

U – Punkt hinter Horopter (ungekreuzte Querdisparation)

K – Punkt vor Horopter (gekreuzte Querdisparation)

Abbildung 12: Blick von hinten auf die Netzhaut und auf die Überlagerung beider Netzhäute bei Abbildung verschiedener Punkte auf den Netzhäuten

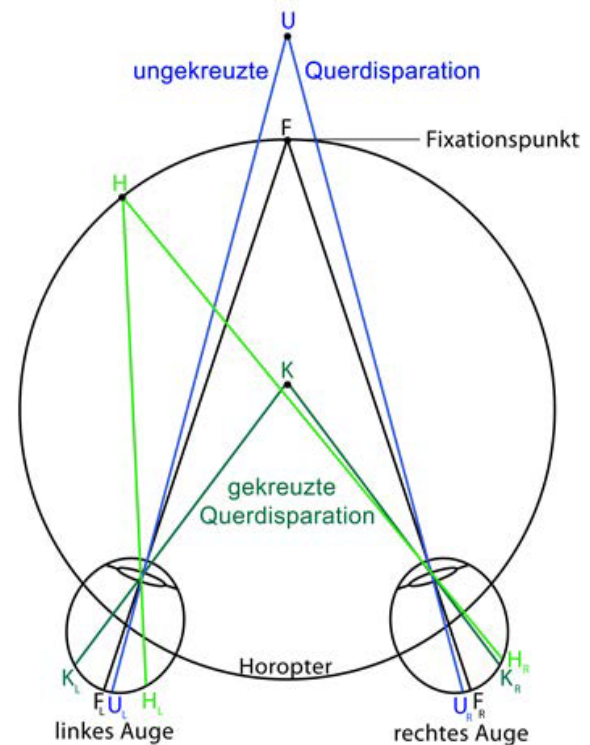


Abbildung 13: Bei Fixation von F treten disparate Netzhautpunkte von K (gekreuzte Querdisparation) und U (ungekreuzte Querdisparation) und korrespondierende Netzhautpunkte von F und H auf [10, Vgl. Abb. S. 41]

¹ Sehrinde

den nichtkorrespondierenden Netzhautpunkten, die unterschiedliche Bildinformationen enthalten. Deshalb werden diese Punkte auch als *disparate Netzhautpunkte* bezeichnet. Dabei wird unterschieden, ob sich ein Objekt vor oder hinter dem Horopter befindet. Wenn das Objekt vor dem Horopter liegt, handelt es sich um eine *gekreuzte Querdissparation*, bei der das Objekt auf den äußeren Randbereichen der Netzhäute abgebildet wird. Bei der ungekreuzten Querdissparation befindet sich das Objekt hinter dem Horopter und wird auf den inneren Randbereichen der Netzhäute abgebildet (Abbildung 12, 13). Aus diesen Entfernungszuordnungen auf bestimmte Netzhautareale kann das visuelle System ableiten, welche Objekte sich vor bzw. hinter dem fixierten Objekt befinden. Durch diese Informationen ist das Gehirn in der Lage ein stereoskopisches Abbild der Umwelt zu erzeugen. Auf neuronaler Ebene gibt es eine Spezialisierung von Nervenzellen im primären visuellen Cortex und den nachfolgenden Verarbeitungsbereichen. So reagierten in Tierversuchen bestimmte Nervenzellen auf die Reizung zweier an gleicher Stelle jedes Auges befindlichen Netzhautpunkte mit unterschiedlichem Bildinhalt. Diese *binokularen Neuronen* reagieren nur bei Reizung beider Netzhautpunkte, was zeigt, dass ihre Funktion mit der Tiefenwahrnehmung zusammenhängen muss [10, S. 39-42].

2.3 Wahrnehmung der Raamtiefe

Im Folgenden soll erläutert werden, welche weiteren visuellen Anhaltspunkte das menschliche visuelle System nutzt, um einen stereoskopischen Raumeindruck zu bekommen. Neben der Querdissparation als wichtigstes Kriterium für räumliche Tiefe, kann das visuelle System auf einige weitere so genannte *Tiefenkriterien der Raumwahrnehmung* (Abbildung 14) zurückgreifen. Diesen Tiefenkriterien liegen verschiedene Eigenschaften zugrunde. So erhält das visuelle System neben *binokularen*¹ Informationen durch die Querdissparation, eine Vielzahl an Informationen durch *monokulare*², *okulomotorische*³ und *bewegungsinduzierte* Kriterien.

Durch die Auswertung der Augenstellung und der Spannung der Augenmuskeln gewinnt der visuelle Apparat Informationen über Entfernungen von fixierten und nicht fixierten Objekten. Damit geben diese als *Konvergenz* und *Akkommodation* bezeichneten okulomotorischen Tiefenkriterien dem visuellen System wichtige Anhaltspunkte für die Tiefenwahrnehmung. Wird ein Objekt betrachtet, so drehen sich die Augen nach Innen und die Blickrichtungen konvergieren bzw. schneiden sich im fixierten Punkt. Der von den Sehachsen eingeschlossene *Konvergenzwinkel* ist bei geringen Entfernungen groß und nimmt mit zunehmender

¹ mit beiden Augen sehend

² mit einem Auge sehend

³ Bewegungen des Auges

EIGENSCHAFT	TIEFENKRITERIUM	ENTFERNUNG		
		0 – 2 m	2 – 30 m	über 30 m
binokular	Querdisparation	✓	✓	
okulomotorisch	Konvergenz	✓		
	Akkommodation	✓		
monokular	Verdeckung	✓	✓	✓
	Relative Größe	✓	✓	✓
	Relative Höhe		✓	✓
	Atmosphärische Perspektive			✓
	Zentralperspektive	✓	✓	✓
	Texturgradient	✓	✓	✓
bewegungsinduziert	Bewegungsparallaxe	✓	✓	
	Zu- und Aufdecken von Flächen	✓	✓	✓

Abbildung 14: Systematik der Tiefenkriterien der Raumwahrnehmung und ihre Anwendung in verschiedenen Entfernungsbereichen [4, Vgl. Abb. S. 191 bzw. 14, Vgl. Tab. 8.1, S. 106]

Entfernung des Fixationspunktes ab. Durch trigonometrische Berechnungen mit Hilfe des Konvergenzwinkels können absolute Objektentfernungen bestimmt werden. Bei sehr großen Objektentfernungen sind die Augen bzw. die Sehachsen parallel zueinander orientiert, wodurch ein Konvergenzwinkel von Null entsteht. Befindet sich das Objekt in Augennähe, entsteht ein großer Konvergenzwinkel. Dabei verdickt sich die Linse wodurch das Objekt fokussiert werden kann. Diese *Akkommodation* ermöglicht dem visuellen System eine Anpassung an die Objektentfernung unter zwei bis drei Metern (darüber ist keine Linsenkrümmung erforderlich), indem die Abbildung auf der Netzhaut durch die Krümmung der Linse scharf gestellt wird. Gleichzeitig nutzt das Gehirn den Akkommodationszustand der Linse als Anhaltspunkt zur Einschätzung der Objektentfernung [10, S. 42-43].

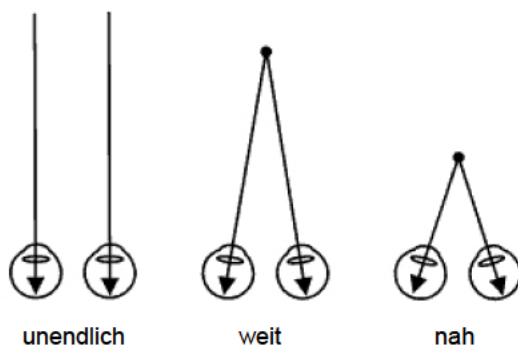


Abbildung 15: unterschiedliche Konvergenzwinkel in Abhängigkeit von der Objektentfernung [10, Abb. S. 43]

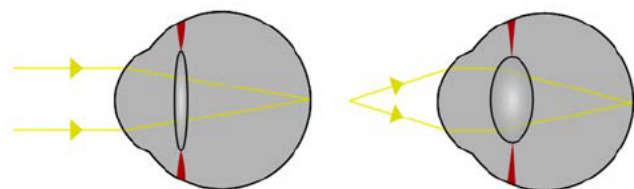


Abbildung 16: Anpassung des Auges an unterschiedliche Objektentfernung durch Verdickung der Linse – Akkommodation [13, Abb. S. 11]

Neben bewegungsinduzierten Tiefenkriterien gibt es noch Tiefenhinweise, die mit einem Auge erfasst und bewertet werden können. Zu diesen monokularen Tiefenkriterien gehören *Verdeckung*, *relative Größe*, *relative Höhe*, *perspektivische Atmosphäre*, *Zentralperspektive* und *Texturgradient*.

Die *Verdeckung* von Objekten ist für das visuelle System ein hinreichender Hinweis für unterschiedliche Entfernungen von Objekten. Wenn ein Objekt durch ein anderes davor platziertes nicht mehr oder nur noch teilweise sichtbar ist, handelt es sich um Verdeckung (Abbildung 17) [4, S. 187]. Durch die Rekonstruktion bzw. Ergänzung der fehlenden Teilstücke des verdeckten Objektes, kann das Gehirn folgern, dass sich das verdeckte Objekt hinter dem anderen befinden muss [5, S. 15]. Das verdeckte Objekt wird als weiter entfernt wahrgenommen als das vordere Objekt, jedoch können durch die Verdeckung keine absoluten sondern nur relative Entfernungen zwischen den betrachteten Objekten abgeschätzt werden [4, S. 187].

Wenn Objekte bekannter Größe unterschiedlich groß erscheinen, folgert das visuelle System daraus, dass sich die Objekte wegen der Größenkonstanz in einer unterschiedlichen Entfernung befinden müssen. Aufgrund dieses Tiefenkriteriums der *relativen Größe* kann der Mensch einschätzen, welches der Objekte sich am dichtesten und welches Objekt bzw. welche Objekte sich weiter entfernt befinden (Abbildung 17) [5, S. 14].

Je höher der tiefste Punkt eines Objektes im Gesichtsfeld liegt, desto weiter weg erscheint das Objekt. Dieser als *relative Höhe* bekannte Tiefenhinweis liefert einen weiteren Anhaltspunkt für die Entfernung eines Objektes (Abbildung 17) [4, S. 187].

Bei der Betrachtung von weiter entfernten Objekten muss das visuelle System durch das Medium Luft hindurch sehen. Da Luft aufgrund von schwebenden Partikeln (z.B. Staub- oder Wasserteilchen) eine Trübheit aufweist, wirken Objekte mit zunehmender Entfernung weniger scharf und kontrastärmer, was als *atmosphärische Perspektive* bezeichnet wird (Abbildung 18) [4, S. 188].

Parallele Linien konvergieren nach den geometrischen Gesetzen der *Zentralperspektive* in der Ferne. Diese Zusammenhänge nutzt das visuelle System ebenfalls zur Konstruktion räumliche Tiefe (Abbildung 17, 19) [10, S. 48].

Treten sich wiederholende, in bestimmten Abständen entfernte, gleiche bzw. ähnliche Grundelemente auf, erscheinen diese mit zunehmender Entfernung dichter gepackt. Dieser als *Texturgradient* bezeichnete Tiefenhinweis steht im Einklang mit dem Kriterium der relativen Größe, d.h. die einzelnen Grundelemente des Texturgradienten erscheinen mit zunehmender Entfernung kleiner [4, S. 189].



Abbildung 17: Verdeckung, relative Größe, relative Höhe, Perspektive [14, Abb. 8.2 S. 100] **Abbildung 18:** atmosphärische Perspektive [15, Abb. S. 232]. **Abbildung 19:** Zentralperspektive [16]

In Abbildung 17 ist die Wirkung mehrerer monokularer Tiefenkriterien besonders anschaulich dargestellt. Drei in der Ebene maßstabsgerechte, identische Abbildungen einer Frau sind so angeordnet und mit einem perspektivischen Hintergrund in Beziehung gesetzt, dass das visuelle System andere Schlussfolgerungen bezüglich der Körpergröße der Frauen im dreidimensionalen Raum zieht. Wenn nur das Kriterium der relativen Größe betrachtet wird, erscheinen die drei Frauen gleich groß. Unter Einbeziehung des perspektivisch dargestellten Hauses erscheinen jedoch die mittlere und die rechte Frau größer als die linke Frau (Die Frau rechts kann auf das Dach sehen). Durch das Kriterium der relativen Höhe in Kombination mit dem Tiefenhinweis der Perspektive erscheint die mittlere Frau größer und weiter entfernt als die linke Frau. Die rechte Frau erscheint größer und weiter entfernt als die anderen Frauen. Trotz der Zweidimensionalität des Bildes versucht das Gehirn, auch bei Betrachtung des Bildes mit nur einem Auge, eine Raumtiefe zu konstruieren.

Wichtige Anhaltspunkte zur Raumwahrnehmung können schließlich durch bewegungsinduzierte Tiefenhinweise, die bei der *Bewegungsparallaxe* und beim *fortschreitenden Zu- und Aufdecken von Flächen* auftreten, gewonnen werden.

Die *Bewegungsparallaxe* ist beim Vorbeibewegen an verschieden entfernten Objekten zu beobachten. Dabei bewegen sich nah gelegene Objekte scheinbar schneller an uns vorbei als weiter entfernte Objekte. Dieser Effekt wird anhand von Abbildung 20 deutlich. Während das Abbild des weit entfernten Hauses beim Vorbeibewegen nur den relativ geringen Netzhautbereich von H_1 bis H_2 überstreicht, wird der Baum hingegen von B_1 bis nach B_2 über einen sehr großen Bereich der Netzhaut abgebildet. Damit Abbildungen nah gelegener Objekte in der gleichen Zeit einen längeren Weg als entfernte Objekte auf der Netzhaut zurücklegen können, müssen sich diese nahen Objekte schneller am Auge vorbei bewegen, als entfernte Objekte [4, S. 190].

Das *fortschreitende Zu- und Aufdecken von Flächen* ist ein weiteres bewegungsinduziertes Tiefenkriterium. Ein Beobachter kann durch eine fortschreitende Bewegung ein entferntes

Objekt durch ein nah gelegenes Objekt zu- oder aufdecken, was den Tiefenhinweis liefert, dass das zu- bzw. aufgedeckte Objekt weiter entfernt als das zu- bzw. aufdeckende Objekt ist (Abbildung 21) [4, S. 190].

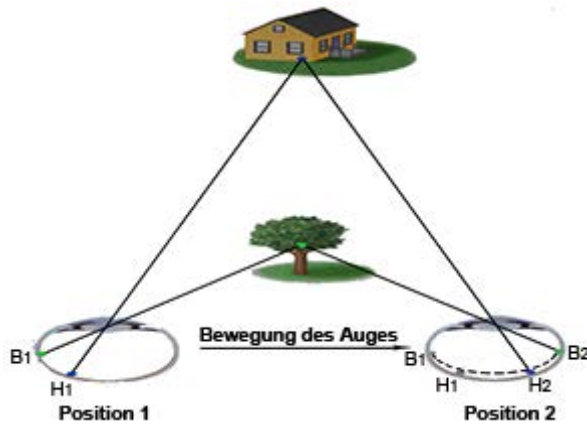


Abbildung 20: Bewegungsparallaxe [4, Vgl. Abb. S. 190]



Abbildung 21: fortschreitendes Zu- und Aufdecken von Flächen [10, Vgl. Abb. S. 45]

2.4 Raumwahrnehmung für die stereoskopische Videoproduktion und die CAVE

Im vorangegangenen Abschnitt wurde deutlich, dass die Querdisparation die wichtigste Komponente für eine räumliche Tiefenwahrnehmung ist. Die Beeinflussung der Querdisparation erfolgt technisch durch die genaue Justage der *Stereobasis*, die den Abstand zwischen beiden Augen bzw. den zur Aufnahme verwendeten beiden Kameraobjektiven darstellt.

Um eine möglichst plausible stereoskopische Wahrnehmung zu erzeugen, muss auch die *Deviation* berücksichtigt werden. Dieses Maß stellt die Differenz zwischen dem Versatz des linken und rechten Teilbildes des entferntesten stereoskopischen Punktes (Fernpunkt) und dem Versatz des linken und rechten Teilbildes des dichtesten stereoskopischen Punktes (Nahpunkt) dar. Es sollte in der Regel 1/30 der gesamten ohne Kopfverdrehen erfassbaren Bildbreite nicht überschreiten [17, S. 1-2]. Neben der Querdisparation sind Akkommodation und Konvergenz als binokulare Tiefenkriterien relevant für die stereoskopische Videoproduktion und die CAVE. Diese beiden Mechanismen sind beim realen Sehen gekoppelt. Während die Augen durch Konvergenz auf den Fixationspunkt gerichtet sind, fokussieren die Augen automatisch auf diesen Punkt (Akkommodation). Dagegen muss der Betrachter diese beiden Mechanismen bei technisch realisierten stereoskopischen Bildern und Filmen in separate Prozesse trennen. Während die Augen auf die Leinwand bzw. den Monitor fokussieren, kann der Konvergenzpunkt bei einer stereoskopischen Szenerie vor oder hinter dem Monitor bzw. der Leinwand liegen. Da der Konvergenzpunkt und der Akkommodationspunkt nicht immer übereinstimmen, muss der Betrachter versuchen beide Prozesse getrennt zu steuern, was jedoch zu Ermüdungen und Kopfschmerzen führen kann [13, S. 15].

Von den monokularen Tiefenkriterien müssen nur einige bei der Planung und Durchführung der stereoskopischen Videoproduktion bzw. bei der Filmwiedergabe in der CAVE berücksichtigt werden. Wegen der geringen Entfernung des Betrachters zur Leinwand von ca. zwei Meter, kann die atmosphärische Perspektive vernachlässigt werden. Ebenso treten keine kleinen, zahlreichen Elemente auf, die als Texturgradient Tiefenhinweise für die Raumwahrnehmung liefern könnten. Hingegen muss bei den mittels Greenscreen-Technik in den Hintergrund eingefügten Konzertsälen im Hinblick auf Verdeckung durch das im Vordergrund befindliche Streichquartett darauf geachtet werden, dass an den Verdeckungsbereichen keine unnatürlichen Effekte wie Ausfransen oder grünliches Leuchten durch die Anwendung des Greenscreen-Verfahrens erzeugt werden. Beim Zusammenfügen des Streichquartetts mit den Konzertsaal-Panoramen müssen die relative Größe, die relative Höhe und die Perspektive beachtet werden. Das Streichquartett muss dabei in der richtigen relativen Größe zur Bühne bzw. zum Konzertsaal dimensioniert und gleichzeitig in der richtigen relativen Höhe platziert werden. Darüber hinaus muss der Streichquartett-Aufsichtswinkel so angepasst werden, dass keine für das visuelle System scheinbare perspektivische Verzerrungen auftreten. Da sich in der CAVE weder Objekte der Projektion noch Versuchspersonen bewegen, sind bewegungsinduzierte Tiefenkriterien hier nicht von Bedeutung.

3 Technische Realisation einer virtuellen Umgebung

Im Folgenden sollen die Grundlagen und Grenzen des visuellen und des auditiven Systems für die menschliche Wahrnehmung anhand ihrer spezifischen Eigenschaften dargestellt werden. Die genaue Kenntnis über diese Eigenschaften mit ihren Möglichkeiten und Grenzen ist eine wichtige Voraussetzung für die technische Erzeugung einer virtuellen Realität. Die Nutzung der daraus resultierenden Erkenntnisse für eine technische Adaption wird in den sich anschließenden Kapiteln erläutert. Unter Berücksichtigung aller Faktoren und technischen Möglichkeiten wird die Realisierung einer CAVE untersucht und anhand der CAVE an der TU Berlin dargestellt.

3.1 Grundlagen und Grenzen der visuellen Wahrnehmung

3.1.1 Sichtbares Licht

Die grundlegende Voraussetzung für visuelle Sinneswahrnehmungen ist die Existenz von Licht. Das sichtbare Licht, das das menschliche Sehen erst ermöglicht, ist nur ein Teil des elektromagnetischen Spektrums, was ein Kontinuum elektromagnetischer Energie darstellt (Abbildung 22). Charakterisiert wird diese Energie durch die Wellenlänge. Der Wertebereich der Wellenlänge erstreckt sich über ca. 16 Zehnerpotenzen, d.h. von 10^{-12} Meter (Gammastrahlung) - 10^4 Meter (Radiowellen). Das menschliche visuelle System kann jedoch nur Wellenlängen zwischen 400 und 700 Nanometern wahrnehmen. Die im sichtbaren Licht vorhandenen Wellenlängen korrelieren mit den vom menschlichen visuellen System wahrnehmbaren Farben. Die Wahrnehmung von Objekten ist nur dadurch möglich, dass das Licht von Objekten direkt in unser Auge reflektiert wird und auf der Netzhaut ein Bild der Objekte erzeugt. Neben diesen Welleneigenschaften besitzt das Licht noch Teilcheneigenschaften, nach denen das Licht in Form von Energieportionen –als Photonen– auftritt. Sowohl die Wellen- als auch die Teilcheneigenschaften des Lichtes spielen eine wichtige Rolle bei der visuellen Wahrnehmung [4, S. 31].

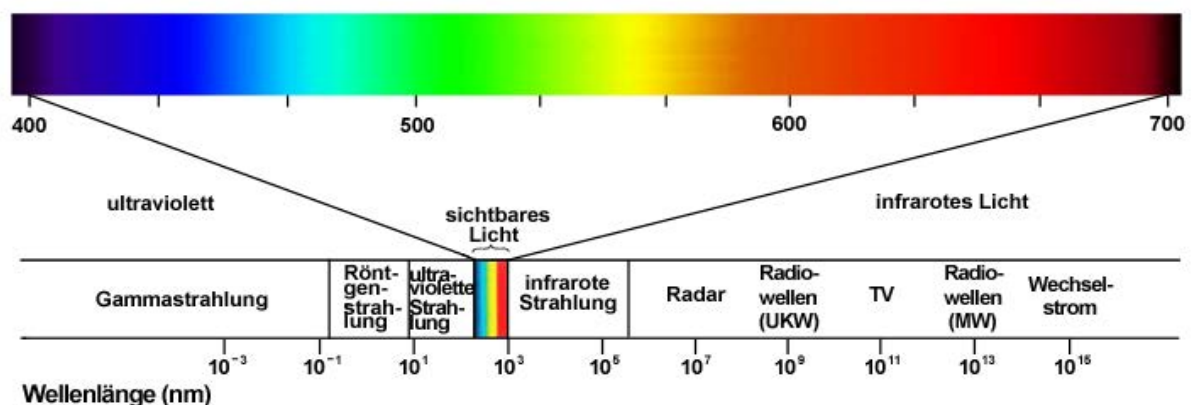


Abbildung 22: Wellenlängen des sichtbaren Lichtes als Teil des elektromagnetischen Spektrums [4, Vgl. Abb. S. 31]

3.1.2 Hellempfindlichkeit und Empfindungsbereich

Wie genau die Augen an bestimmte Wellenlängen angepasst sind, wird bei der Betrachtung von *spektralen Hellempfindlichkeitskurven* deutlich. Aus Abbildung 23 geht hervor, dass die Empfindlichkeit der Augen generell zu den mittleren sichtbaren Wellenlängen hin, welche überwiegend Farben zwischen dunkelgrün und gelb entsprechen, am größten ist. Der Unterschied zwischen Stäbchen- und Zapfensehen zeigt sich in einer höheren Empfindlichkeit der Stäbchen für kurzwelliges Licht (Maximum bei 500nm) gegenüber der zu relativ langwelligem Licht (Maximum 560nm) verschobenen Empfindlichkeit der Zapfen. Daraus resultiert eine höher werdende Empfindlichkeit für kurzwelliges Licht zum grün-blau gehenden Spektrum hin bei der Adaption an Dunkelheit, die mit einer Verlagerung des Sehens von den Zapfen zu den Stäbchen einhergeht. Diese Farbempfindlichkeitsverschiebung wird nach ihrem Entdecker als *Purkinje Effekt* bezeichnet (Abbildung 23) [4, S. 43].

Die größere Lichtempfindlichkeit von Stäbchen gegenüber Zapfen hat mehrere Ursachen. Zum einen wird weniger Licht benötigt, um eine Reaktion eines Stäbchenrezeptors im Vergleich zu einem Zapfenrezeptor auszulösen. Des Weiteren werden die Antworteinheiten mehrerer, mit einer Ganglienzelle verschalteten, Stäbchen aufsummiert und können dadurch die Empfindlichkeitsschwelle der Ganglienzelle eher überschreiten als die jeweils nur mit einer Ganglienzelle verschalteten Zapfen (Abbildung 24). Aus dieser für die hohe Lichtempfindlichkeit optimierten Verschaltung der Stäbchen resultiert ein geringeres räumliches Auflösungsvermögen bzw. eine geringere Sehschärfe gegenüber der Verschaltung der Zapfen. Werden z.B. entsprechend Abbildung 24 die äußeren Zapfen bzw. Stäbchen von je einem Lichtpunkt stimuliert, so liefert nur eine Stäbchenganglienzelle, aber beide äußeren Zapfenganglienzellen eine Antwort. Da die Zapfenganglienzellen örtlich getrennt sind, werden die zugehörigen Netzhautpunkte der Zapfen getrennt registriert, was einer höheren örtlichen Auflösung entspricht [4, S. 49-51].

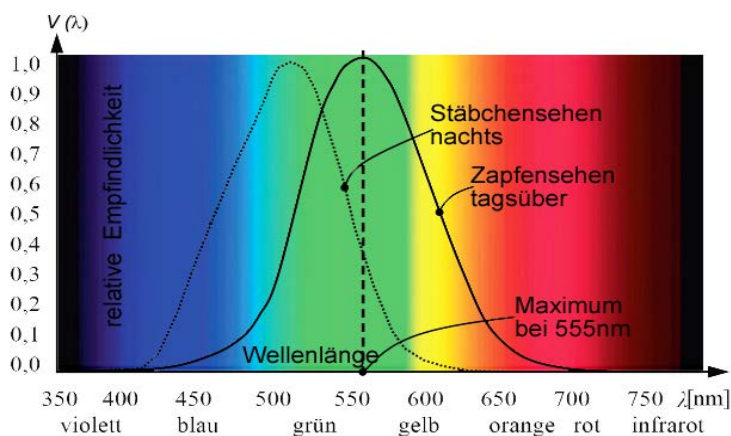


Abbildung 23: spektrale Hellempfindlichkeitskurven für das Stäbchen- und Zapfensehen [18, Vgl. Abb. S. 4]

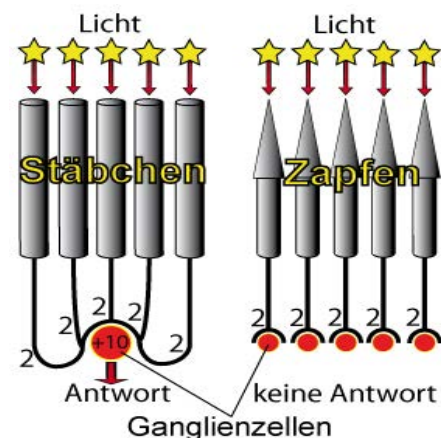
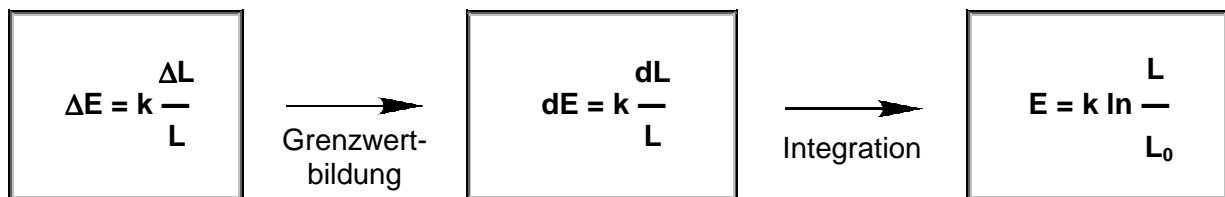


Abbildung 24: Verschaltung von Stäbchen und Zapfen [4, Vgl. Abb. S. 49]

Das menschliche visuelle System ist nach einer gewissen Zeit in der Lage seine Helligkeitsempfindlichkeit in einem sehr weiten Bereich von ca. 8 Zehnerpotenzen an die Umgebungsleuchtdichte anzupassen. Der genaue Zusammenhang zwischen Reizstärke und dem resultierenden Empfindungsunterschied lässt sich über einen weiten Bereich der Reizstärke durch das nichtlineare, logarithmische *Weber-Fechner-Gesetz* beschreiben. Es besagt, dass die Änderung der Helligkeitsempfindung ΔE proportional zum Verhältnis aus Leuchtdichteänderung ΔL und der Leuchtdichte L ist [19, S. 20-21]:



E – Helligkeitsempfindung
 ΔE – Änderung der Helligkeitsempfindung
 L – Leuchtdichte

ΔL – Leuchtdichteänderung
 L_0 – Leuchtdichte-Wahrnehmungsschwelle
 k – Proportionalitätskonstante

[20, S. 4-6]

Entsprechend der logarithmischen Kennlinie (Abbildung 25) werden niedrige Lichtreize nahe der Leuchtdichte-Wahrnehmungsschwelle $L = L_0$ (entspricht R_0 in Abbildung 25) verstärkt und so „gut empfindbar“ gemacht. Dagegen werden sehr große Lichtreize stark abgeschwächt. Dadurch wird das visuelle System vor Überlastung geschützt und kann hohe Helligkeitswerte in der Regel problemlos wahrnehmen [20, S. 6]. Das bedeutet, dass der Mensch *relative Helligkeitsunterschiede* sehr gut wahrnehmen kann [21, S. 40]. Die Helligkeitsempfindung hängt jedoch noch von anderen Faktoren ab. Sie wird u. a. von angrenzenden oder umgebenden Arealen beeinflusst. Dieser als Simultankontrast bezeichnete und hier auf die Helligkeit bezogene Effekt kann anhand Abbildung 26 nachvollzogen werden. Obgleich der fast geschlossene Ring nur einen Grauwert besitzt, wirkt das linke offene Ende des Ringes heller als das rechte offene Ende des Ringes. Offenbar beeinflusst der umgebende Hintergrund die Helligkeitsempfindung. Die retinalen Zellen auf die der helle Hintergrund fällt, werden stark stimuliert, sodass diese die Zellen im Bereich der rechten Ringhälfte stark hemmen. Dieser Effekt der so genannten *lateralen Inhibition* (seitliche Hemmung) bewirkt andererseits, dass die Zellen auf die der schwarze Hintergrund fällt die Zellen der eingeschlossenen linken Ringhälfte nur schwach hemmen. Da die Zellen unter der linken Ringhälfte weniger von den umgebenden Zellen in ihrer Aktivität gehemmt werden, wird ihre Antwort weniger reduziert als die Antwort der unter der rechten Ringhälfte befindlichen Zellen, die somit stark reduziert wird. Diese Zellen feuern weniger, wodurch das offene Ende der rechten Ringhälfte dunkler wirkt als das offene Ende der linken Ringhälfte. Nur an der unteren geschlossenen Ringhälfte ist dieser Effekt nicht zu beobachten. Das

bedeutet, dass man das Phänomen nicht allein mit lateraler Inhibition erklären kann [4, S. 56-57].

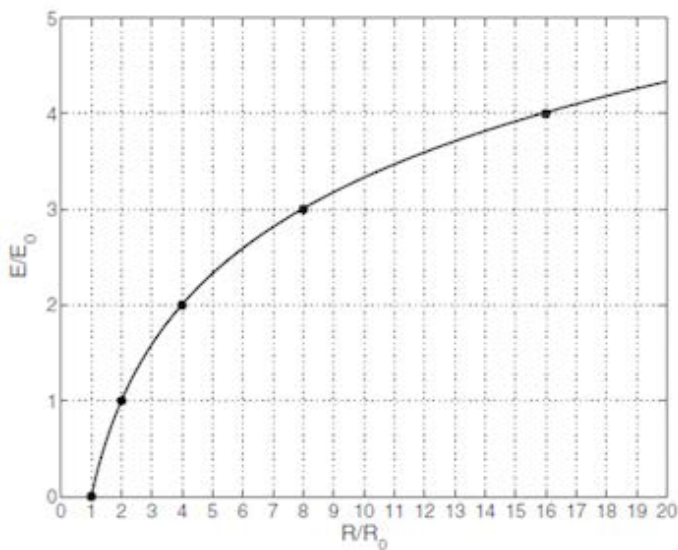


Abbildung 25: Qualitativer Zusammenhang zwischen Reiz R und Empfindung E [20, Abb. S. 5]

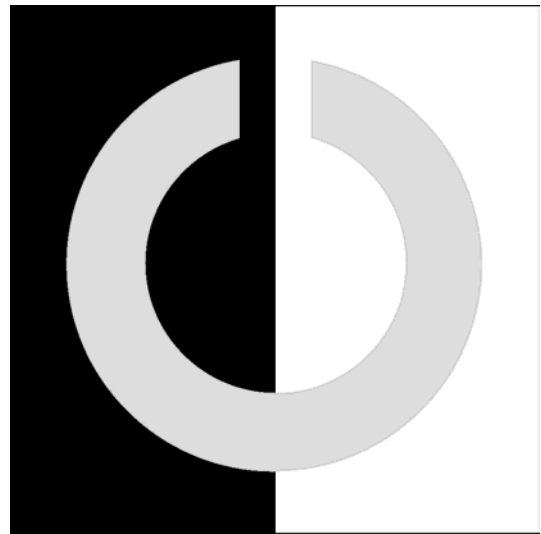


Abbildung 26: simultaner Helligkeitskontrast

3.1.3 Farbwahrnehmung

Wie schon in Abschnitt 3.1.1 beschrieben, ermöglichen die Zapfen bei ausreichender Leuchtdichte die Wahrnehmung von Farben, die mit den Wellenlängen des sichtbaren Lichtes korrelieren.

Durch Untersuchungen der Netzhaut wurde festgestellt, dass sich die vorwiegend in der Fovea centralis lokalisierten Zapfen in drei verschiedene Gruppen einteilen lassen, die für jeweils eine Farbe eine besonders hohe Empfindlichkeit aufweisen. Dem entsprechend besitzen die S-Zapfen¹ für blaues Licht, die M-Zapfen² für grünes Licht und die L-Zapfen³ für rotes Licht die jeweils höchste Empfindlichkeit. Jeder der drei Zelltypen reagiert jedoch nicht nur auf eine Farbe, sondern auf einen bestimmten Bereich des Farbspektrums (Abbildung 27). Während sich die S-Zapfen auf einen eher schmalen Bereich beziehen, decken die L-Zapfen das breiteste Spektrum ab. Die Reizantworten der verschiedenen Zapfen-Typen werden schon auf der retinalen Ebene vorverarbeitet [22, S. 81-82]. Diese Kombination verschiedener Reizantworten basiert auf dem Prinzip der additiven Mischung, bei der aus den drei Farbwerten Rot, Grün und Blau die entsprechenden Mischfarben erzeugt werden. Beispielsweise wird die Farbe Gelb aus der Mischung der Reizantworten roter und grüner Zapfengruppen wahrgenommen [23, S. 69]. Dieser additive Mischprozess der verschiedenen

¹ S – kurze Wellenlänge; engl. short wavelength

² M – mittlere Wellenlänge; engl. medium wavelength

³ L – lange Wellenlänge; engl. long wavelength

Farbanteile erfolgt über mehrere Zwischenschritte. Während zunächst aus Rot und Blau ein gemischter Farbanteil erzeugt wird, ergibt sich aus Grün und Blau ein weiterer gemischter Farbanteil. Dann wird aus diesen beiden Mischfarbtönen der Blau-Gelb-Anteil berechnet. Parallel dazu findet die Berechnung eines Mischfarbtönen aus Rot und Grün statt. Nach einer weiteren additiven Mischung dieser beiden Farbanteile ergibt sich daraus die endgültige Farbtonmischung, die schließlich an das Gehirn weitergeleitet wird (Abbildung 28). Parallel zu der Farbtonmischung kann durch die Analyse der drei Farbanteile die Helligkeit ermittelt werden, die ebenfalls an das Gehirn weitergeleitet wird. Die in Abbildung 28 grau dargestellten, helligkeitsempfindlichen Stäbchen erreichen bei Tageslicht ihre Sättigung (hellstes Weiß) und liefern so am Tage keine Helligkeitsdifferenzen mehr. Durch diese gesamte retinale Vorverarbeitung wird die notwendige Rechenleistung zur Weiterverarbeitung im Gehirn deutlich verringert [24].

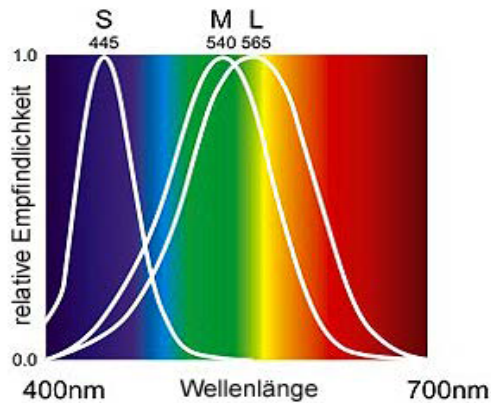


Abbildung 27: relative Empfindlichkeit der drei Zapfentypen in Abhängigkeit von der Wellenlänge [25]

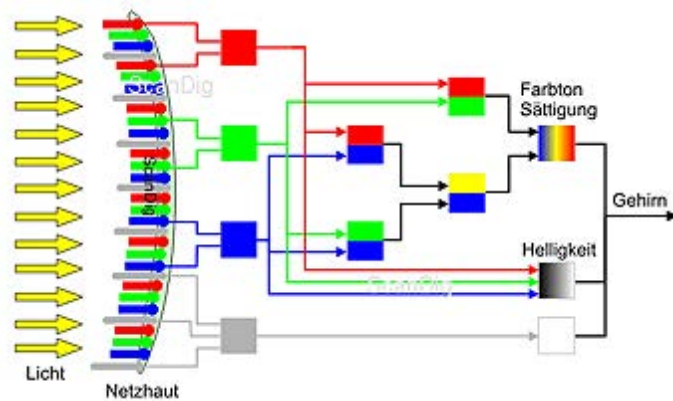


Abbildung 28: Erfassung und Vorverarbeitung des Lichtreizes durch Frequenz- bzw. Farbanalyse mittels der drei Zapfentypen bei Tageslicht [24]

Mittels der Farbtonmischung aus den drei Zapfentypen ist es dem visuellen System möglich ca. 10 Millionen Farben zu unterscheiden. Darüber hinaus ist es in der Lage sich durch *chromatische Adaption* innerhalb von maximal 30 Minuten an Beleuchtungsquellen mit Licht unterschiedlicher Wellenlängen bzw. an eine unterschiedliche so genannte Farbtemperatur des Lichtes anzupassen [26, S. 91]. Diese Farbtemperatur ist eine äquivalente, quantitative Darstellung des Spektrums einer Lichtquelle, die die relativen Anteile der unterschiedlichen Wellenlängen in dieser Lichtquelle widerspiegelt. Damit ist es möglich, den Farbton von Lichtquellen gleicher Intensität auf einer linearen Skala (in Kelvin) abzubilden (Abbildung 29). Die Farbtemperatur entspricht der Temperatur eines erwärmten schwarzen Körpers, bei der dieser schwarze Körper Licht der gleichen Farbe wie die Lichtquelle abgibt [27]. Trotz Objektbeleuchtungen mit unterschiedlich spektral zusammengesetzten Lichtquellen ver-

schiedener Intensitäten liefert das visuelle System des Menschen konstante Farbeindrücke. Diese für die Objekterkennung wichtige Eigenschaft des visuellen Systems wird durch die relative Konstanz im Verhältnis der Rezeptorantworten ermöglicht. Daraus resultiert die *Farbkonstanz* in der menschlichen Wahrnehmung [14, S. 82]



Abbildung 29: Einordnung natürlicher und künstlicher Lichtquellen [28, Tab. S. 217] auf der Skala der Farbtemperaturen [29, Abb. S. 101]

3.1.4 Sensumotorik

Die visuelle Wahrnehmung kann als sensumotorischer Prozess betrachtet werden. Der Seheindruck entsteht dabei aus der Zusammensetzung der Augenbewegungen und den Lichtreizungen der Netzhaut. Forschungen mit retinal stabilisierten Bildern ergaben, dass die Blickbewegungen des Auges für das bewusste Sehen notwendig sind. Insbesondere werden auch bei Fixationen immer -wenn auch nur sehr kleine- Augenbewegungen ausgeführt, die offenbar für die visuelle Wahrnehmung notwendig sind. Im Gegensatz zu technischen Systemen mit stabilen Abbildungen auf Leinwänden bzw. Bildschirmen treten daher keine permanent stabilen Bilder auf der Netzhaut auf. Die wichtigsten sensumotorischen Leistungen des Sehens werden durch Akkommodation und Konvergenzbewegungen der Augen erbracht. Bei der Akkommodation wird die Fokussierung von unterschiedlich entfernten Objekten auf der Netzhaut durch Steuerung der Ziliarmuskeln, die die Linse wölben, ermöglicht. Die Fixation eines Punktes wird durch die Eindrehung beider Augen nach Innen erreicht. Durch diese Konvergenzbewegungen ist die Erfassung von Objekten in unterschiedlicher Entfernung möglich. Je näher sich das betrachtete Objekt am Auge befindet, desto größer ist der dabei entstehende Winkel. Aus diesem Konvergenzwinkel und der Auswertung des Akkommodationszustandes der Linse kann das Netzhaut-Bild im Gehirn zu einem scharfen stereoskopischen Eindruck des fokussierten Objektes verrechnet werden (Vgl. Abschnitt 2.3, S. 11-12). Dieser resultierende Seheindruck eines Objektes fungiert in der Regel als Steuergröße für die Ausführung weiterer motorischer Augenbewegungen. Unbewusst werden während des Sehvorganges ständig Ziliarmuskeln für die Akkommodation kontrahiert, der Konvergenzwinkel angepasst sowie kleine Blickverlagerungen bei der Fixation ausgeführt [22, S. 91-93]. Um Objekte genauer erfassen zu können, führen die Augen schnelle Bewegungen, so genannte Sakkaden, aus. Mittels kamerabasierter Eye-tracker, die die aktuelle Position der Pupille und damit den Blickwinkel des Auges ermitteln,

lassen sich Blickbewegungsverläufe genauer untersuchen (Abbildung 30). In diesem Zusammenhang wurde festgestellt, dass Versuchspersonen Objekte durch eine Folge sich abwechselnder Fixationen und Sakkaden betrachteten (Abbildung 31). Typischerweise werden bei der Betrachtung eines Objektes bzw. einer Szenerie drei Fixationen in einer Sekunde ausgeführt [4, S. 133-134]. Das Ineinandergreifen von motorischer Verarbeitung, in Form von Augenbewegungen, und die sensorische Verarbeitung, in Form der Reizung lichtempfindlicher Fotorezeptoren, können als zusammenwirkende Prozesse des Sehens verstanden werden [22, S. 92].



Abbildung 30: Eyetracker; dient der Erfassung von Blickbewegungsverläufen [30]

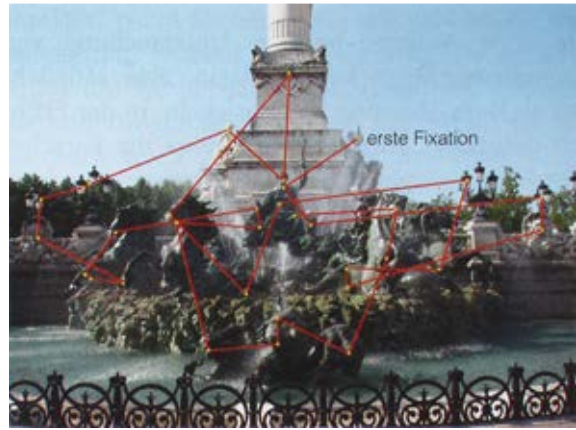


Abbildung 31: individueller Blickbewegungsverlauf einer Person beim Betrachten eines Bildes
gelbe Punkte: Fixationen
rote Linien: Sakkaden [4, Abb. S. 134]

3.1.5 Örtliches Auflösungsvermögen

Die höchste Farbunterscheidungsfähigkeit und die beste örtliche Auflösung der Netzhaut werden in der Fovea centralis, die nur einen sehr geringen Teil der Netzhaut einnimmt, erreicht (Vgl. Abschnitt 2.1, S. 8). Damit möglichst viele fokussierte, hoch aufgelöste Farbbildinformationen eines Objektes mittels der vielen Zapfen in der Fovea centralis erfasst und an das Gehirn weitergeleitet werden können, müssen möglichst viele Bereiche des Objektes in der Fovea centralis abgebildet werden. Dies wird durch die schnelle Abfolge von Fixationen und Sakkaden zur Objektbetrachtung möglich [22, S. 78-79]. Trotz der ausschließlichen Erfassung von scharfen Farbbildern in der Fovea centralis kann der Mensch einen großen Bereich der Umwelt mit einem Blick visuell wahrnehmen. Dieses durch beide aber mindestens durch ein Auge erfassbare Gesichtsfeld, das ein Mensch bei einer festen Kopf- und Augenstellung wahrnehmen kann, beträgt entsprechend perimetrischen¹ (Gesichtsfeld-) Messungen (Abbildung 32) horizontal ca. 180° und vertikal ca. 130°. Aufgrund der unterschiedlichen Verteilung bzw. Dichte von Stäbchen und Zapfen ist das Gesichtsfeld für farbige Reize kleiner als für Hell / Dunkel Reize. Durch die Überlagerung der Gesichtsfelder beider Augen ist in einem Bereich von ca. 120° (horizontal) binokulares

Sehen möglich (Abbildung 33). Dadurch kann das visuelle System nach Weiterverarbeitung der Bilder Tiefe wahrnehmen und Entfernungen bzw. Abstände genau bestimmen [31, S. 76-77].

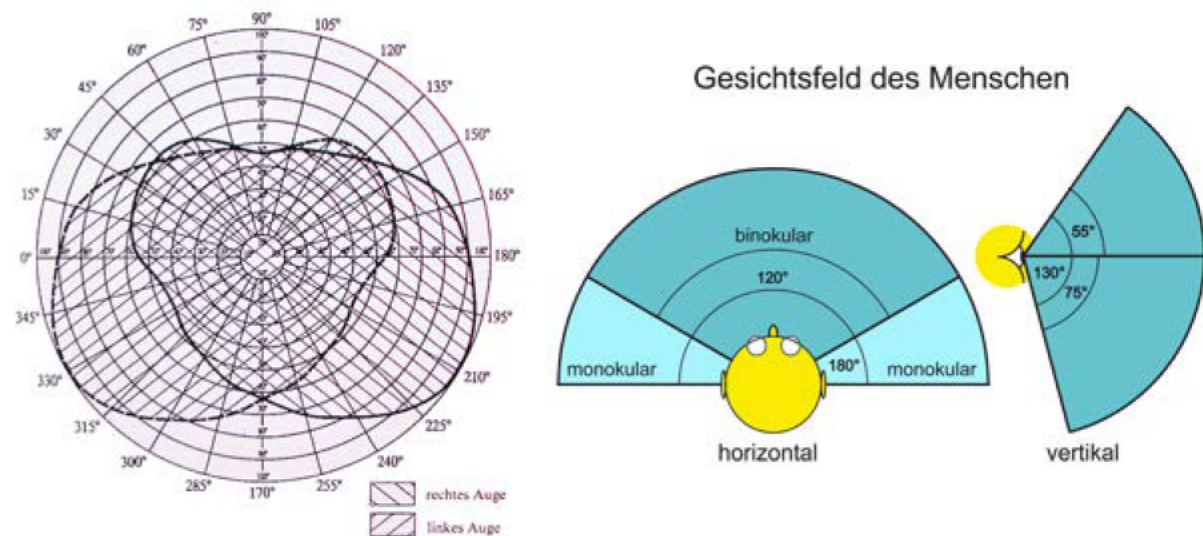


Abbildung 32: Gesichtsfeld des Menschen (perimetrische Messung) [31, Abb. S. 77]

Abbildung 33: horizontales und vertikales Gesichtsfeld des Menschen (Daten entsprechend Abbildung 32)

Das kleine Netzhautareal der scharfen Farbbilderfassung in der Fovea centralis erzeugt dagegen sowohl vertikal als auch horizontal im Gesichtsfeld des Menschen nur einen kleinen scharf abgebildeten Bereich von $1,5^\circ$ [32, S. 6]. Das bedeutet, dass die Wahrnehmung feiner Details und damit ein hohes Auflösungsvermögen nur im Bereich der Fovea centralis, also im direkten Blickmittelpunkt des Gesichtsfeldes gewährleistet wird.

Das örtliche Auflösungsvermögen gibt an, bei welchem Abstand zwei Punkte oder Linien gerade noch als getrennt wahrgenommen, also aufgelöst werden können. Je geringer dieser Abstand ist, desto höher ist die, auch als Sehschärfe bekannte, örtliche Auflösung. Um die Auflösung unabhängig vom Objektstand zu dimensionieren, wird als Maß der Sehschärfe ein Winkel definiert. Dieser Winkel befindet sich zwischen den beiden Strahlen, die von den Messpunkten/-linien ausgehend sich im Knotenpunkt des Auges schneiden [33].

Um das Auflösungsvermögen zu quantifizieren, wird untersucht, wie viele Schwarz/Weiß-Linienpaare im Blickfeld bezogen auf ein Grad noch als getrennt wahrgenommen werden können. In Abbildung 34 ist der Zusammenhang zwischen der menschlichen Wahrnehmung von Helligkeitsunterschieden bzw. Kontrasten und der Raumfrequenz dargestellt. Dabei ist zu erkennen, dass mit zunehmender, auch als Ortsfrequenz bezeichneter, Liniendichte, die Unterscheidbarkeit abnimmt. Die Grenze bei der noch separate Linien wahrgenommen werden können, liegt bei ca. $30 \text{ LP}/^\circ$ (Linienpaaren je Grad), d. h. **60 Linien / 1 Grad**. Daraus ergibt sich:

bei **1 Grad = 60 Winkelminuten**

60 Linien / 60 Winkelminuten

1 Linie / 1 Winkelminute

Das bedeutet, dass die Differenzierung einzelner Linien bis in den Bereich von einer Winkelminute möglich ist. Darüber hinaus ist erkennbar, dass Kontraste am deutlichsten zwischen 2 und 10 Linienpaaren sichtbar werden.

Der Kurvenverlauf in Abbildung 34 kann als Modulationstransferfunktion (MTF) behandelt werden. Diese Funktion gibt das Verhältnis der Modulation m zur maximalen Modulation m_{\max} bei verschiedenen Liniendichten an:

$$\text{MTF} = \frac{m}{m_{\max}} \quad \text{mit} \quad m = \frac{s_{\max} - s_{\min}}{s_{\max} + s_{\min}} \quad \begin{array}{l} s_{\max} - \text{maximale Amplitude} \\ s_{\min} - \text{minimale Amplitude} \end{array}$$

Die Modulationstransferfunktion wird ebenfalls als Maß für technische Systeme (z.B. für Objektive) verwendet. Die Kontrastempfindlichkeit und das Auflösungsvermögen hängen des Weiteren von der Beleuchtungsstärke und vom Modulationsgrad der Bildvorlage ab [23, S. 22-24].

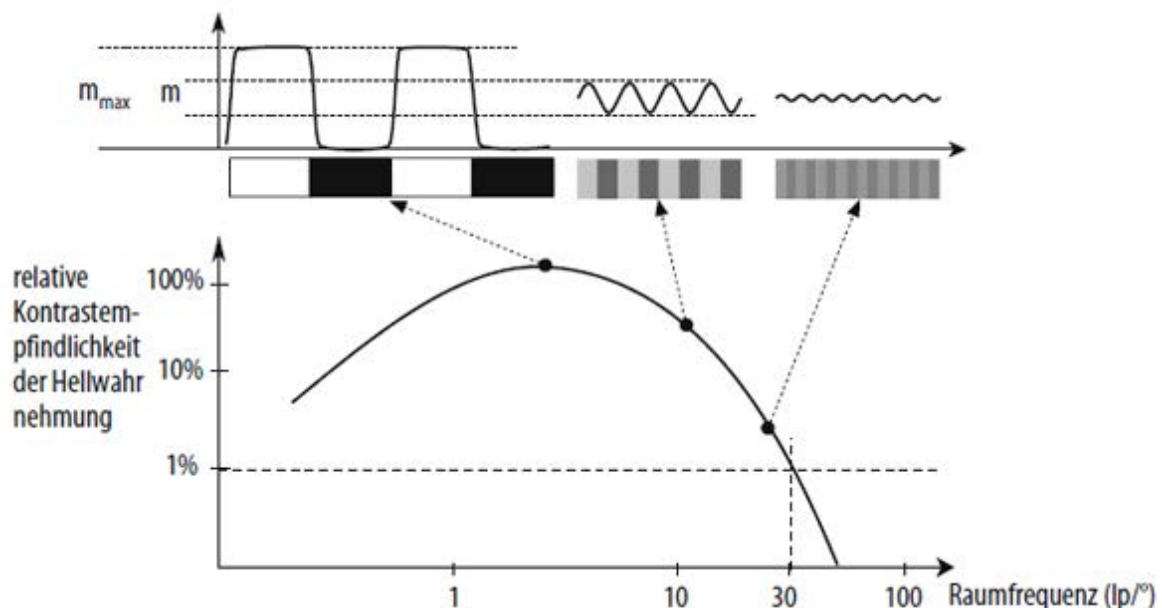


Abbildung 34: Örtliches Auflösungsvermögen des Auges im Zusammenhang mit der MTF [23, Abb. S. 23]

3.1.6 Zeitliches Auflösungsvermögen

Neben dem örtlichen besitzt das menschliche Auge ein spezifisches zeitliches Auflösungsvermögen, das angibt, in welcher Zeiteinheit zwei Lichtreize noch als getrennt wahrgenommen werden können. Je größer diese Reizwiederholfrequenz wird, desto weniger können die Lichtreize bzw. Bilder getrennt werden. Ab einer bestimmten, von der Lichtintensität, der Wellenlänge und der Objektgeschwindigkeit abhängigen, *Flimmererschmelzungsfrequenz* verbinden sich die einzelnen Lichtreize bzw. Bilder zu einer

kontinuierlichen flimmerfreien Beleuchtung bzw. zu kontinuierlichen flimmerfreien Objektbewegungen. Aus der Abhängigkeit von der Lichtintensität resultieren unterschiedliche Beträge der Flimmerverschmelzungsfrequenzen. So liegt diese zeitliche Auflösungsgrenze im Tageslicht bei ca. 50 Hz und im Dämmerlicht bei ca. 10 Hz [33].

Im Gegensatz zur örtlichen Auflösung, bei der die Trennung zweier Bildpunkte für eine hohe Auflösung bzw. Schärfe das Ziel ist, wird bei der zeitlichen Auflösung keine Trennung sondern eine Verschmelzung der zeitlichen Reize angestrebt, damit u.a. Videosequenzen kontinuierlich und nicht als Folge von Einzelbildern wahrgenommen werden können.

Das Orts- und das Zeitverhalten des visuellen Systems hängen eng zusammen. So verringert sich das scharfe Sehen und damit das örtliche Auflösungsvermögen bei schnellen zeitlichen Änderungen von bewegten Objekten bzw. Bildern. Ebenso verringert sich das zeitliche Auflösungsvermögen bei einer geringeren örtlichen Leuchtdichteverteilung [19, S. 29-30]. Das bedeutet u.a., dass dunklere Filme mit relativ geringer Leuchtdichte bei niedriger Bildwiederholfrequenz ohne Flimmern wahrnehmbar sind, während bei Filmen mit hoher Leuchtdichte bei gleicher Bildwiederholfrequenz bereits ein Flimmern auftreten kann.

3.2 Adaption der Technik an die visuelle Wahrnehmung

Um eine immersive, visuelle Wahrnehmung zu erreichen, muss ein optisches Übertragungssystem zum Einsatz kommen, das sich optimal an die im vorangegangenen Abschnitt dargestellten Eigenschaften der visuellen menschlichen Wahrnehmung anpasst. Zunächst sollte ein geeignetes Aufnahmesystem dimensioniert und verwendet werden, dessen Funktionen und Parameter zu denen des Auges äquivalent sind. Da die Verwendung der CAVE zeitunabhängig erfolgen soll, muss sowohl das Aufnahme- als auch das Wiedergabesystem eine Speicherung der Filmsequenzen ermöglichen. Diese sollen unkomprimiert und redundant in einem geeigneten digitalen Datenformat vorliegen und flexibel abrufbar sein.

3.2.1 Optische Aufnahmesysteme und Parameter

Obwohl das Auge als Teil der belebten Natur ein Produkt der Evolution ist, können viele Funktionen mittels geeigneter Technik adaptiert werden. Im Folgenden werden technische Systeme und deren Funktionen in Bezug zum Auge beschrieben und verglichen.

Das Aufnahmesystem, das zu der Funktion des menschlichen Auges die zurzeit wohl größte Äquivalenz darstellt, ist ein digitaler Camcorder bzw. Digitalkamera mit Filmaufnahmefunktion. Trotz einiger wesentlicher Unterschiede zwischen Auge und Kamera kann durch Einsatz eines modernen Camcorders als Aufnahmesystem eine gute Qualität des Filmmaterials zur Verwendung in der CAVE erreicht werden.

¹ SI-Basiseinheit der Leuchtdichte

Im grundlegenden Aufbau von Auge und Kamera gibt es viele Gemeinsamkeiten, was auch aus der Vorbildfunktion des Auges bei der Entwicklung der Kamera resultiert. So bestehen beide Systeme aus einer Linsen- Projektionsflächen-Kombination, die mittels Lichteinfall durch eine Blende bei der Kamera bzw. durch die Pupille beim Auge auf dem Bildsensor bzw. der Netzhaut ein umgekehrtes, verkleinertes, seitenverkehrtes Bild entstehen lässt (Abbildung 35).

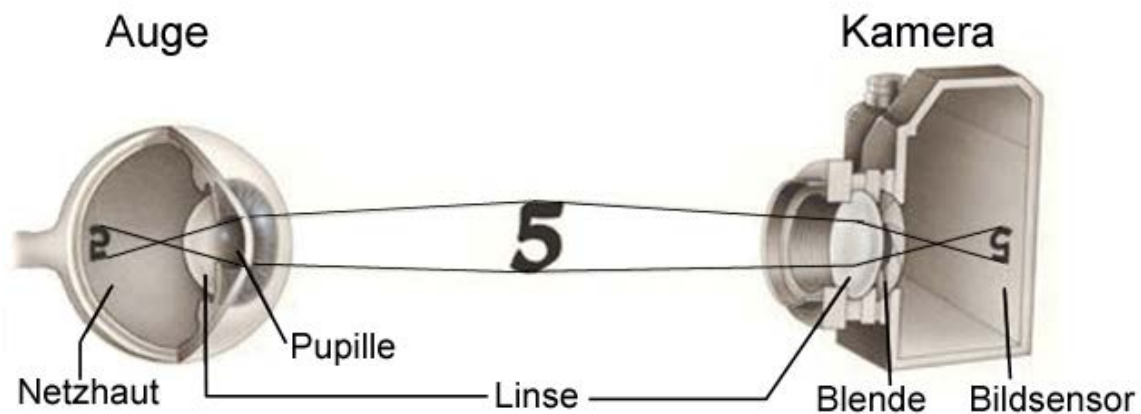


Abbildung 35: Strahlengang und Bildentstehung bei Auge und Kamera [34, Vgl. Abb.]

Eine wichtige Voraussetzung bei der technischen Umsetzung eines optischen Aufnahmesystems ist eine optimale Bildentstehung auf der Projektionsfläche. Das wird durch die Verwendung von lichtempfindlichen den menschlichen Fotorezeptoren in Funktion ähnlichen Sensoren bzw. Materialien, die das Spektrum des sichtbaren Lichtes verarbeiten können, erreicht.

Während bei belichtetem analogem Filmmaterial (lichtempfindliches Silberhalogenid) eine chemische Reaktion und anschließende chemische Interaktion mit verschiedenen Farbstoffen zur Entwicklung des Filmmaterials und damit zur Bildentstehung führt, wird bei der sich weiter verbreitenden, elektronischen, digitalen Bilderfassung mittels *Bildsensor* Silizium bzw. SiO_2 als lichtempfindliches Material verwendet. Durch den *lichtelektrischen Effekt* können Photonen aus dem Silizium Elektronen herauslösen, die dann als Ladungsmengen registriert und in digitaler Form gespeichert werden können. Die Erfassung der Ladungsmengen aller matrixartig angeordneten Bildpunkte auf dem Bildsensor ermöglicht die digitale Speicherung und Weiterverarbeitung der Bilddaten und das spätere Bearbeiten bzw. Ausdrucken digitaler Bilder [35]. Das für elektronische Bildsensoren verwendete lichtempfindliche Silizium besitzt eine spektrale Empfindlichkeit von ca. 350-1000 nm und umfasst damit den für das menschliche Auge relevanten Wellenlängenbereich des sichtbaren Lichtes von ca. 380-720 nm. Darüber hinaus kann der Wellenbereich des Infrarots (720-1000 nm) erfasst und sichtbar gemacht werden [36].

Als Bildsensoren kommen zurzeit hauptsächlich *CCD-Sensoren*¹ und *CMOS-Sensoren*² zum Einsatz. CCD-Sensoren bestehen aus tausenden kleinen lichtempfindlichen Pixeln³, die matrixartig auf einem rechteckigen Silizium-Träger angeordnet sind. Während der Aufnahme eines Bildes entstehen je nach Objekthelligkeit unterschiedlich große Ladungsmengen in den einzelnen Pixeln. Diese Ladungsmengen werden jedoch nicht direkt ausgelesen, sondern wandern zunächst Zeile für Zeile über die Pixelmatrix bis sie in das Ausleseregister am Sensorrand gelangen. Von dort werden die Ladungsmengen in den Ausleseverstärker verschoben, wo sich nach der Signalverstärkung die Umwandlung der unterschiedlichen Pixel-Ladungsmengen durch einen Analog-Digital-Wandler in unterschiedliche digitale Spannungswerte vollzieht [37]. Der zu jedem Bildpunkt ermittelte Spannungswert stellt dabei die Helligkeitsinformation des entsprechenden Pixels dar. Der gesamte Bereich zwischen dem geringsten und dem größten Helligkeitswert gibt den so genannten Dynamikumfang des Sensors an. Diese wichtige Kenngröße kann Grauwertabstufungen von 8-Bit (256 Abstufungen) bis 16-Bit (65536 Abstufungen) beinhalten [36]. Dieser sensorischen Leistung entspricht beim Auge die Funktion der Stäbchen, die nur Helligkeitswerte unterscheiden können. Während die Wahrnehmung von Farbe beim Auge mit Hilfe der Zapfen erfolgt, gibt es beim CCD-Sensor mehrere Möglichkeiten der Farberzeugung.

Am häufigsten zur Anwendung kommt die so genannte *Bayer-Matrix*, die als Filter vor den Bildsensor positioniert wird (Abbildung 36). Dabei sind 50 % der Filterelemente mit einer grünen, 25% mit einer blauen und 25 % mit einer roten Schicht belegt. Diese unterschiedliche Gewichtung resultiert aus den Empfindlichkeitseigenschaften des Auges, nach denen die für grün empfindlichen L-Zapfen das breiteste Spektrum abdecken (Vgl. Abschnitt 3.1.3, S. 20). Die Pixel unter den blauen Filterelementen erfassen den Blauanteil, die grünen den Grün- und die roten den Rotanteil des vom aufzunehmenden Objekt einfallenden Lichtes. Aus den drei entstandenen, schachbrettartigen Teilbildern interpoliert der Prozessor der Kamera mit Hilfe entsprechender Softwarealgorithmen ein farbiges Gesamtbild. Dabei werden die Informationen der drei Farbanteile jeweils zu einem dreifarbigem Pixel zusammengerechnet. Durch elektronische Scharfzeichnungsfilter wird die durch Interpolation entstandene Weichzeichnung korrigiert und ein Weißabgleich (Anpassung an die Farbtemperatur der Lichtquelle) zwischen den drei Teilfarben durchgeführt. Nach allen notwendigen Bildberechnungen wird das fertige Bild im RAW-Format (unkomprimiert) oder im JPG-Format (komprimiert) auf einer Speicherkarte abgelegt. Da die Algorithmen zu den Bildberechnungen nicht standardisiert sondern kamera- bzw. herstellerabhängig sind, gelten diese neben der Optik und Mechanik der Kamera als wichtiges Qualitätskriterium für die Bildqualität einer Kamera. Um die Qualität zu verbessern, entwickelte die Firma Fuji aus der quadratischen

¹ Charge Coupled Device - ladungsgekoppeltes Bauelement

² Complementary MetalOxide Semiconductor – Metalloxid-Halbleiter

³ Bildpunkte

Bayer-Matrix einen Sensor mit achteckigen Elementen. Dieser so genannte Super-CCD-Chip erfasst laut Fuji feinere Strukturen und erzielt damit eine bessere Auflösung. Um die Interpolation der Teilbilder zu vermeiden, liegen beim X3-Chip von Foveon die drei Farbsensoren bzw. die Farbfilter übereinander. Dadurch erhält jedes Pixel direkt die gesamte Farbinformation. Damit kann eine höhere Auflösung und Bildqualität als mit der Bayer-Matrix erreicht werden [28, S. 293-294]

Eine interpolationsfreie Bilderfassung ist ebenfalls mit der 3CCD Technik möglich (Abbildung 37). Für jede Grundfarbe steht dabei je ein Bildsensor zur Verfügung. Das einfallende Licht wird durch einen aus Prismen zusammen gesetzten, so genannten Farbteiler in die entsprechenden Rot-, Grün- und Blauanteile separiert. Jedes Teilspektrum trifft auf jeweils einen Bildsensor. Ohne Interpositionsverluste kann das Bild anschließend aus den drei Teilbildern errechnet werden. Durch die gleichzeitige Belichtung der einzelnen Bildsensoren ist die Erzeugung eines farbigen Livebildes mit beweglichen Motiven möglich. Aufgrund dessen und wegen der hohen Kosten für drei Bildsensoren werden 3CCD-Sensoren vorwiegend in hochwertigen Camcordern bzw. Kameras eingesetzt [38].

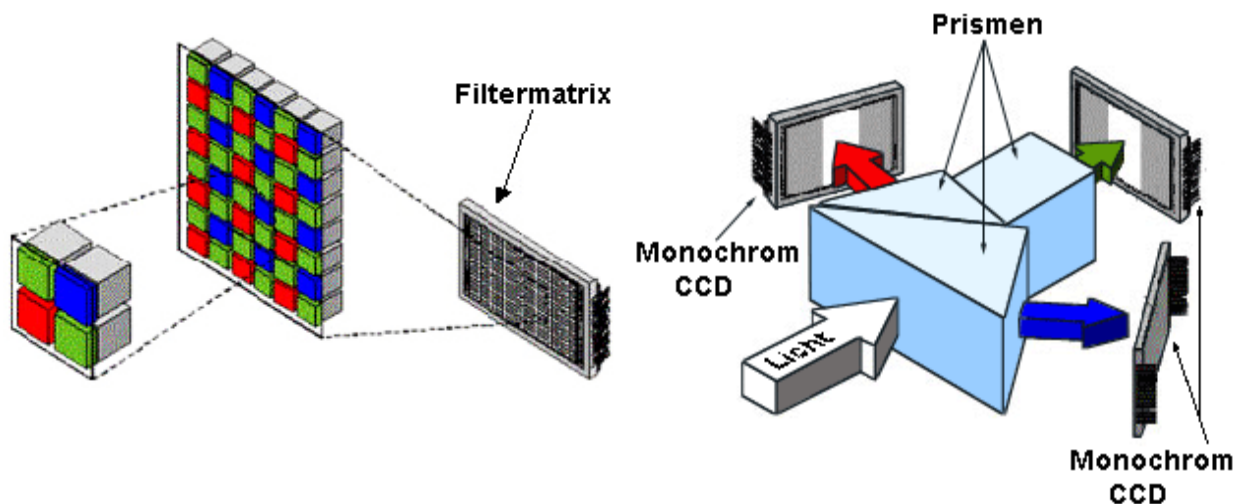


Abbildung 36: Bayer-Matrix – Filtermaske für CCDs mit einem Bildsensor [39, Abb.]. **Abbildung 37:** 3CCD Technik [38, Vgl. Abb.].

Eine andere Technologie basiert auf dem Einsatz hochintegrierter CMOS-Schaltkreise. Jeder lichtempfindliche Bereich ist dabei mit einem bzw. bei moderneren so genannten Aktiv Pixel Sensoren (APS) mit mindestens drei Transistoren verbunden. Je nach einfallender Lichtintensität wird eine entsprechende Menge an Ladungsträgern gebildet. Im Gegensatz zu CCD-Sensoren können die Ladungen auf den CMOS-Sensoren direkt verarbeitet bzw. die einzelnen Pixel mit Hilfe einer xy-Matrix direkt adressiert werden, was u.a. eine schnellere Bildfolge und ein Betreiben des Sensors sowohl im Interlaced Mode (Erzeugung von Halbbildern) oder Progressiv Mode (Erzeugung von Vollbildern) ermöglicht. Durch Transistoren werden die Bildinformationen über direkte Leitungen aus dem Sensor geführt

und müssen nicht über den gesamten Chip transportiert werden. Bei den zurzeit verwendeten APS werden bis zu fünf Transistoren eingesetzt, die eine direkte Adressierung, Verstärkung und Ableitung der Ladung bis hin zu einer aufwändigen Signalverarbeitung direkt auf dem Chip und so eine unmittelbare Ausgabe eines Digitalsignals ermöglichen. Durch diesen Aufbau bedingt, ergibt sich bei CMOS-Chips insbesondere im Vergleich zur CCD-Technik ein hoher Signal /Rauschabstand und ein sehr geringer Leistungsbedarf, der im mW-Bereich liegt [23, S. 370-371].

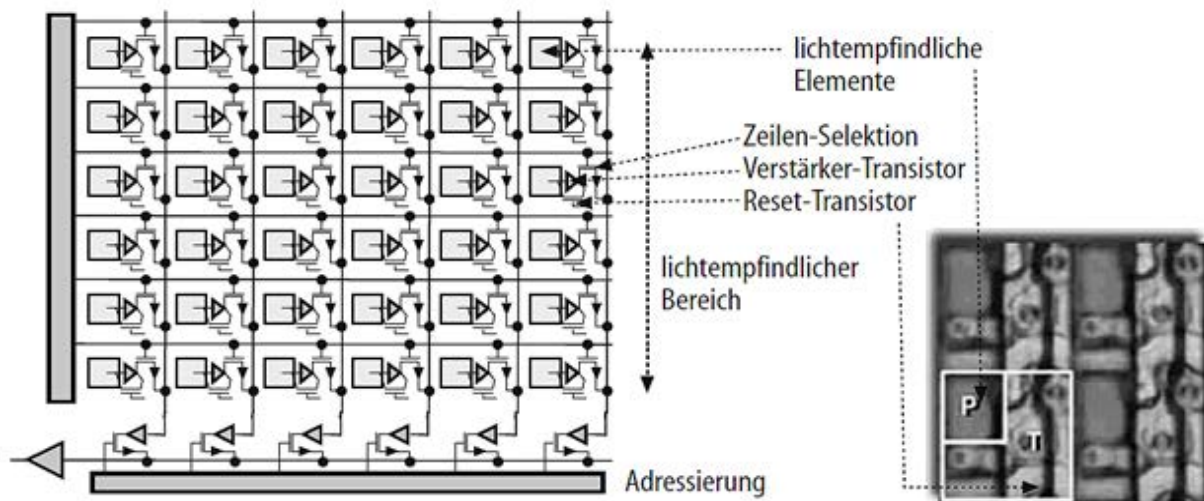


Abbildung 38: Funktionsprinzip eines CMOS- Sensors; rechts: Ausschnitt aus der Sensorfläche [23, Abb. S. 371]

Während die auf der Netzhaut in der Fovea centralis abgebildeten Objekte nur in einen Bereich von $1,5^\circ$ des Gesichtsfeldes scharf und farbig wahrgenommen werden, liefern die vorgestellten Bildsensoren auf der kompletten Sensorfläche die gleiche Auflösung und in jedem Pixel eine gleich gewichtete Helligkeitsinformation. Da das visuelle System durch die schnelle Abfolge von Fixationen und Sakkaden jedoch zahlreiche visuelle Informationen in kurzer Zeit über die Umwelt sammelt und im Gehirn weiterverarbeitet, kann es die geringe Schärfe bzw. Auflösung und die abnehmende Farbwahrnehmung außerhalb der Fovea centralis kompensieren.

Zurzeit verfügen Kamerasysteme über Bildsensoren mit *Auflösungen* bis zu 50 MP. Canon stellte 2010 einen CMOS-Sensor mit 120 MP vor, dessen Auflösung die des menschlichen Auges bei weitem übertrifft [40]. Dennoch können im professionellen Bereich auch Bildsensoren mit wesentlich geringeren Auflösungen zur Anwendung kommen. Die 18 MP Kamera Canon 7D besitzt beispielsweise ein Pixelauflösungsvermögen auf dem Sensor von $9,4 \mu\text{m}$. Damit besitzt sie die achtfache Auflösung des Auges, die bei ca. $80 \mu\text{m}$ liegt und bietet somit aufnahmeseitig die Voraussetzung für eine plausible visuelle Wahrnehmung [41, S.228].

Neben der Auflösung ist das *Sensorformat* ein wichtiger Parameter des Bildsensors. Während kleinere Sensoren häufig in kleinen Kameras und Handys verwendet werden, kommen in Spiegelreflex und hochwertigen Sucherkameras das APS-C - (ca. $16 \times 25\text{mm}$)

oder das Kleinbildformat (24 x 36mm; auch als Vollformat bezeichnet) zum Einsatz. Je größer die Sensorfläche ist, desto geringer ist das Rauschen des Sensors. Der Bildwinkel und der Preis steigen hingegen mit zunehmender Sensorfläche [42]

Sehr flexibel erfolgt beim Auge die Regelung des Lichteinfalls, der Schärfe und des Weißabgleichs. Während sich diese Parameter beim Auge in der Regel automatisch einstellen, müssen diese bei hochwertigen Kamera- und Camcordersystemen meist manuell justiert werden. Diese flexiblen Reaktionen des Auges resultieren aus der physiologischen Zusammensetzung und der neuronalen Steuerung.

So besteht die Augenlinse hauptsächlich aus Wasser und speziellen Proteinen, zu deren Eigenschaften Transparenz und Elastizität gehören [43, S. 5]. Während beim Auge die Akkommodation durch Krümmung der Linse realisiert wird (Vgl. Abschnitt 2.3, S. 12), ist dies bei aus Glas bestehenden Kameralinsen nicht möglich. Durch die festgelegte Wölbung der Linse ist auch die Brennweite konstant. Die Justierung der Schärfe hängt außer von der Brennweite noch von der Gegenstandsweite und der Bildweite ab. Die Zusammenhänge dieser optischen Größen sind in der so genannten *Linsengleichung* zusammengefasst:

$\frac{1}{f} = \frac{1}{g} + \frac{1}{b}$	f = Brennweite - Abstand zwischen Brennpunkt und Linsenmittelpunkt b = Bildweite - Abstand zwischen Linse und Bildsensor / Netzhaut g = Gegenstandsweite - Abstand zwischen Linse und Gegenstand
---	---

Aus der Gleichung geht für Kamerasysteme hervor, dass bei konstanter Brennweite **f** der Kameralinse und veränderlicher Gegenstandsweite **g** (z.B. bewegter Gegenstand) der Abstand der Linse vom Bildsensor (Bildweite **b**) angepasst werden muss, damit die Gleichung erfüllt wird und die Fokussierung des Gegenstandes erhalten bleibt (Abbildung 39).

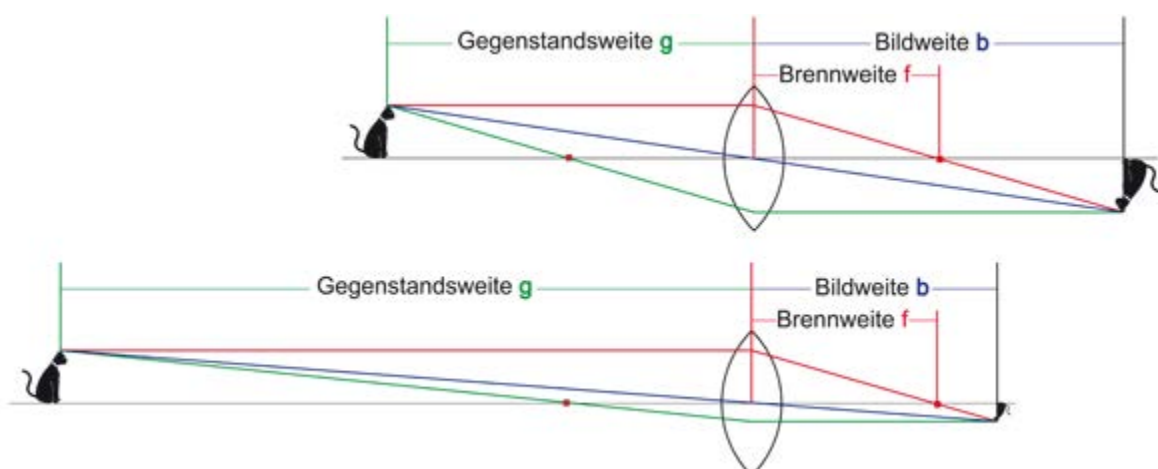


Abbildung 39: Strahlengang und Bildentstehung durch Sammellinse mit konstanter Brennweite [44, Vgl. Abb.]
oben: Gegenstand nah an Linse; unten: Gegenstand weit entfernt von Linse

Dagegen ist bei Anwendung der Gleichung auf das Auge der Abstand der Linse zur Netzhaut (Bildweite b) konstant. Die Anpassung an die Änderung der Entfernung eines Objektes (Gegenstandsweite g) kann beim Auge nur durch eine Änderung der Brennweite f mittels Krümmen der Linse (Akkommodation) realisiert werden [44].

Um bei Kamerasystemen hohe Lichtstärken zu erreichen, optische Fehler korrigieren und unterschiedliche Brennweiten realisieren zu können, sind mehre, auf einer optischen Achse rotationssymmetrisch angeordnete Linsen zu einem *Objektiv* zusammengefasst. Durch den Einsatz konvexer und konkaver Linsen aus verschiedenartigem Glas ist beispielsweise die Korrektur optischer Fehler möglich. Aus der Summe aller Einzelbrechkräfte, die sich aus positiv bewerteten Sammellinsen und negativ bewerteten Zerstreuungslinsen zusammensetzen, ergibt sich die Gesamtbrechkraft eines Objektivs. Abgesehen von Zoomobjektiven mit variabler Brennweite wird zwischen Weitwinkel-, Normal- und Teleobjektiven unterschieden. Von der Brennweite hängen viele Parameter ab: Bildgröße, Abbildungsmaßstab, Bildwinkel, Entfernung zwischen Kamera und Gegenstand, Schärfentiefebereich und die Raumwirkung. Insbesondere bestimmt der Bildwinkel den Bildausschnitt der Szenerie. Während bei Objektiven mit kleiner Brennweite ein großer Bildausschnitt mit einem großen Bildwinkel entsteht, bildet ein Objektiv mit großer Brennweite einen kleinen, aber dafür größer erscheinenden Bereich einer Szene mit kleinem Bildwinkel ab (Abbildung 40) [28, S.194-195].

Beim menschlichen Auge sind mehrere Komponenten an der Lichtbrechung beteiligt. Neben der Linse, die insbesondere die Variabilität der Brennweite und damit die Akkommodation ermöglicht, besitzt die Hornhaut die größte Brechkraft. Zusammen mit dem zwischen Hornhaut und Linse befindlichen Kammerwasser, der Linse und dem Glaskörper ergibt sich ein Objektiv-ähnliches optisches System. Weit entfernte Objekte werden bei ungekrümmter Linse (Ziliarmuskel entspannt) mit maximaler Brennweite auf der Netzhaut scharf abgebildet. Die Brennweite entspricht dann in etwa der Bildweite, die beim Auge ca. 23 mm beträgt. Bei maximaler Linsenkrümmung durch Ziliarmuskelanspannung wird eine minimale Brennweite von ca. 19 mm erreicht [45, S. 51-52]. Objektive von Kamerasystemen besitzen dagegen je nach Anwendung meist feste Brennweiten von 8 mm (Fischaugenobjektiv) bis hin zu 500 mm bei Teleobjektiven.

Ebenso flexibel wird beim Auge die durch die Pupille einfallende Lichtmenge mittels der sich schnell schließenden Iris automatisch geregelt und somit die Netzhaut vor zu großer Lichteinstrahlung geschützt (Vgl. Abschnitt 2.1, S. 7). Im Bereich von Kamerasystemen werden der Funktion der Iris entsprechende *Blenden* mit manuell veränderlichem Durchmesser eingesetzt, die den nutzbaren Bereich der Linse und damit den auf den Bildsensor fallenden Lichtstrom regulieren. Des Weiteren können dadurch linsenbedingte Abbildungsfehler durch Ausblenden fehlerbehafteter Randstrahlen verringert werden [23, S. 386].

Die Blende wird durch die Größe der Blendenzahl repräsentiert. Da diese von der Objektivbrennweite und dem Durchmesser der Blendenöffnung abhängt, gilt folgende Gleichung:

$k = \frac{f}{d}$	k = Blendenzahl f = Brennweite des Objektivs d = Durchmesser der Blendenöffnung
-------------------	--

Aus der Gleichung ergibt sich das umgekehrte Verhältnis von Blendenzahl **k** zu Blendenöffnung **d**. International sind die Blendenzahlen in einer Blendenreihe genormt. Beginnend mit 1 erhöht sich die Blendenzahl jeweils um den Faktor 1,4 ($\sqrt{2}$) [28, S.196].

Zusammen mit der Blende bestimmt die *Belichtungszeit* -die Dauer der Belichtung- die auf den Bildsensor bzw. die Netzhaut auftreffenden Lichtmenge. Um bei unterschiedlichen Blenden die gleiche Lichtmenge zu erhalten, korrespondieren verschiedene Belichtungszeiten mit jeweils einer komplementären Blendenzahl zu einem Belichtungsprodukt mit jeweils gleicher Lichtmenge [46, S. 18].

Ein wichtiger Parameter bei der Dimensionierung von Bild- bzw. Filmaufnahmen ist die *Schärfentiefe*. Dieses Maß gibt an, in welchem Bereich zweier Ebenen senkrecht zur optischen Achse alle Objekte scharf abgebildet bzw. scharf wahrgenommen werden. Eine scharfe Abbildung wird nicht nur in einem Punkt, sondern im Bereich des so genannten Zerstreuungskreises erreicht. Der noch scharf abgebildete Punkt eines Zerstreuungskreises, der der Linse am nächste liegt, wird als Nahpunkt bezeichnet, der noch scharf abgebildete Punkt eines Zerstreuungskreises, der dem aufzunehmenden Objekt am nächsten ist, heißt Fernpunkt. Sowohl beim Auge als auch bei Kameras hängt die Schärfentiefe von der Brennweite, der Blende und von der Entfernung des aufzunehmenden bzw. wahrzunehmenden Objektes ab. Die Schärfentiefe erhöht sich je kürzer die Brennweite, je größer der Aufnahmeabstand und je kleiner der Durchmesser der Blendenöffnung (Abbildung 41) ist [28, S. 197].

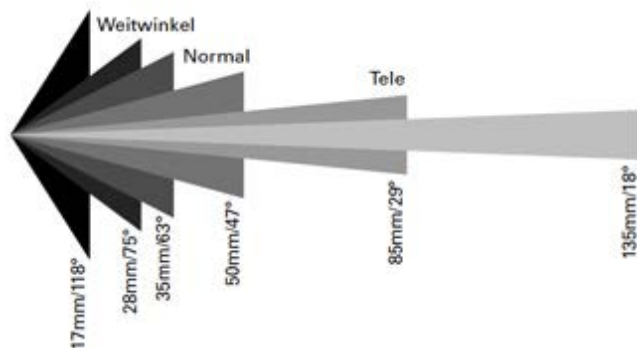


Abbildung 40: Bildwinkel und Brennweiten unterschiedlicher Objektivarten [28, Abb. S.196]

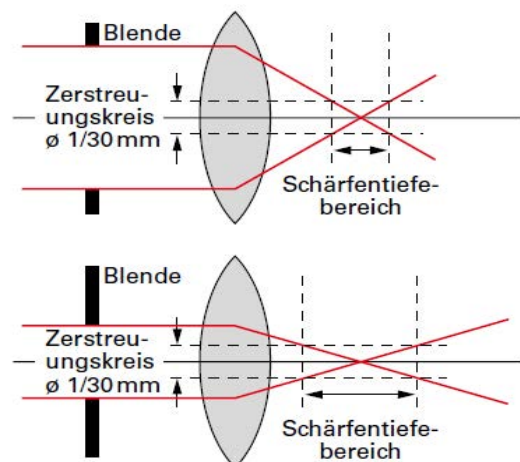


Abbildung 41: Schärfentiefebereich in Abhängigkeit von der Blende; oben: große Blendenöffnung → kleine Schärfentiefe unten: kleine Blendenöffnung → große Schärfentiefe [28, Abb. S.197]

Die Anpassung an unterschiedliche Farbtemperaturen, die beim visuellen System automatisch erfolgt (Vgl. Abschnitt 3.1.3, S. 21-22), wird im Kamerabereich mittels *Weißabgleich* realisiert. Neben einem automatischen Weißabgleich, der sogar bei professionellen Kameras im EB-Einsatz¹ zur Anwendung kommen kann, wird bei den meisten professionellen Produktionen und wissenschaftlichen Projekten ein manueller Weißabgleich durchgeführt, da sich Automatikparameter auf Standardsituationen beziehen und deshalb die Möglichkeit einer Fehljustierung nicht ausgeschlossen werden kann. Dies gilt ebenso für die automatische Blendensteuerung und die automatische Fokussierung [23, S. 408-409]. Ein manueller Weißabgleich wird mittels einer zum Vergleich aufzunehmenden Grau- oder Weißkarte durchgeführt, indem die drei Farbanteile Rot, Grün und Blau so gewichtet werden, dass sie einen neutralen Weiß- bzw. Grauwert ohne Farbstich ergeben [28, S. 219].

Während die Pixeldichte auf dem Sensor (Sensorauflösung) größer als die örtliche Auflösung des Auges sein sollte, muss die Bildwiederholfrequenz größer als die Flimmerverschmelzungsfrequenz –die zeitliche Auflösung des Auges– sein. Das ist die Frequenz, bei der zwei aufeinander folgende Lichtreize nicht mehr als getrennt wahrgenommen werden können, sondern flimmerfrei zu einem kontinuierlichen Film verschmelzen. Sie liegt im Tageslicht bei ca. 50 Hz (Vgl. Abschnitt 3.1.6, S. 26). Dazu wird bei der technischen Umsetzung (z.B. Kino) jedes der 24 Bilder zweimal pro Sekunde projiziert. Die durch diese Frequenzverdopplung entstandene Bildwiederholfrequenz von 48 Hz genügt, um bei den im Kino verwendeten geringen Leuchtdichten von ca. 50 cd/m² (Candela pro Quadratmeter¹) einen Film weitgehend flimmerfrei betrachten zu können. [23, S. 25]. Im Fernseh- und Videobereich liegen die Bildwiederholfrequenzen bei 50 Hz (PAL-Fernsehnorm, 2 x 25 Halbbilder/s) bzw. 60 Hz (NTSC-Fernsehnorm, 2 x 30 Halbbilder/s) [28, S. 981].

Wie im vorangegangenen Abschnitt dargestellt, sind der prinzipielle Aufbau, die grundlegenden Funktionen und Parameter von Auge und Camcorder- bzw. Kamerasystemen sehr ähnlich, obwohl sich Auge und Kamera teilweise sehr stark in der Art und Weise der einzelnen Funktionen unterscheiden. So setzt sich beispielsweise das scharfe Farbsehen beim Auge aus der Kombination der nur einen kleinen Bildwinkel abbildenden Fovea centralis und den umgebungserfassenden Sakkaden zusammen, welche im Gehirn berechnet, bewertet und zu einem Gesamtbild zusammengefügt werden. Dagegen erfolgt bei Kamerasystemen eine gesamte, gleichmäßig gewichtete Erfassung einer Szene und deren unbewertete Weiterverarbeitung und Speicherung im System. Das erschwert mitunter den direkten Vergleich zwischen dem biologischen System der Augen und den technologischen Systemen von Camcordern und Kameras.

¹ Elektronische Berichterstattung im Bereich der Fernsehproduktion

Die wichtigsten Funktionen und Parameter beider Systeme sind in folgender Übersicht zusammengefasst:

FUNKTION / PARAMETER	AUGE	KAMERA / CAMCORDER
Farberfassung	mittels Zapfen, hauptsächlich in der Fovea centralis	mittels Farbfilter, z.B. Bayer-Matrix
Helligkeitsanpassung	automatisch durch Irisanpassung	meist manuell durch Blendenanpassung
örtliche Auflösung	eine Winkelminute, 80 μm auf Retina	9,4 μm auf CMOS-Sensor bei 18 MP
zeitliche Auflösung	ca. 50 Hz	48 Hz (Kino), 50 Hz (TV), größere zeitliche Auflösungen möglich
Fokussierung	automatisch und schnell durch Akkommodation: Änderung der Linsenwölbung durch Ziliarmuskel	meist manuell durch Abstandsänderung von Linse zu Bildsensor
Weißabgleich	automatisch	meist manuell
Bildwinkel	horizontale: 180° vertikal: 130°, aber: scharfer Farbbereich nur in der Fovea centralis	mit speziellen Objektiven horizontal: 180° vertikal: 180° möglich
Bildwinkeländerung	nur durch Kopfdrehung/-neigung möglich, kein Zoomen möglich	mittels Zoomobjektiv möglich
Brennweite	19 mm - 23 mm	je nach Anwendung: mit Weitwinkel-, Normal- und Teleobjektiven: 8 mm - 500mm möglich
Schwenkbewegungen des Aufnahmesystems	durch Sakkaden sind sehr schnelle Schwenkbewegungen möglich	relativ langsame Schwenkbewegungen mittels Schwenkarm
Stereoskopische Bilderfassung	durch Augenabstand und neuronale Weiterverarbeitung im Gehirn möglich	mit stereoskopischen Kameras und Camcordern möglich (Abschnitt 3.2.3)

Abbildung 42: Vergleich von Auge und Kamerasystemen in Bezug auf wichtige Funktionen und Parameter

3.2.2 Optische Wiedergabesysteme und Parameter

Nachdem die grundlegenden Eigenschaften und Parameter eines optischen Aufnahmesystems erörtert wurden, sollen nun optische Systeme zur Wiedergabe der aufgenommenen Bilddaten insbesondere im Hinblick auf Eignung für die CAVE vorgestellt werden. Während das optische Aufnahmesystem eine technische Adaption zum Auge darstellt, muss ein optisches Wiedergabesystem die Eigenschaft besitzen, dem Auge eine der Umwelt gleichende Szenerie darzubieten, die eine plausible Wahrnehmung ermöglicht.

Wiedergabesysteme haben die Aufgabe die aufgenommenen Bild- bzw. Filmdaten in sichtbares Licht umzuwandeln, um vom visuellen System wahrgenommen werden zu können. Während so genannte aktive Wiedergabewandler selbst Licht emittieren, funktionieren passive Wiedergabewandler mittels elektronisch gesteuerter Lichttransmission oder -reflexion. Obwohl im 20. Jh. die Kathodenstrahlröhre am weitesten verbreitet war, setzt sich seit Beginn des 21. Jh. die Flachbildtechnologie durch. Diese besitzt gegenüber der Röhrentechnik zahlreiche Vorteile wie geringeres Volumen, geringeres Gewicht und

geringerer Energiebedarf. Neben diesen Eigenschaften sollten Wiedergabewandler folgende Forderungen erfüllen: große Helligkeit, geringe Reaktionszeiten, großer Betrachtungswinkel, hohes Auflösungsvermögen und große Abbildungsfläche. Im Heimbereich werden Bild-diagonalen von einem Meter und mehr erreicht. Bei speziellen Großbildprojektionen (wie der SEACEN-CAVE) sind dagegen sehr viel größere Abbildungen möglich [23, S. 450]. Die Systematik von Wiedergabewandlern ist in Abbildung 43 dargestellt.

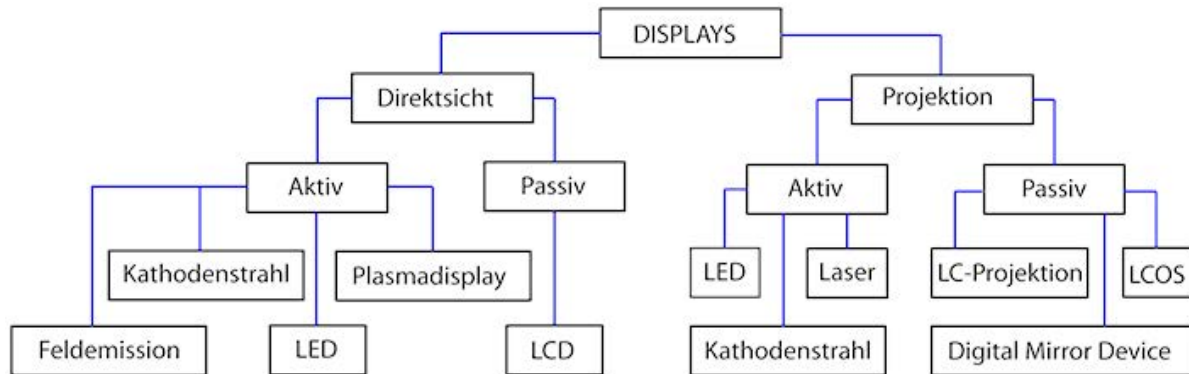


Abbildung 43: Übersicht über Bildwiedergabewandler [23, Vgl. Abb. S. 450]

Aufgrund der zahlreichen Wiedergabesysteme werden im Folgenden nur die für die SEACEN-CAVE relevanten Bildwiedergabewandler vorgestellt. Da Direktsichtgeräte derzeit nicht in der Größe einer CAVE erhältlich bzw. zu teuer sind, muss ein Projektionssystem zum Einsatz kommen. Die Anforderungen des SEACEN-Projektes beinhalten, dass die erzeugten stereoskopischen Panoramen auf einer großen Leinwand mit hoher Auflösung und passenden Helligkeits- und Kontrastverhältnissen plausibel wiedergegeben werden und die Anschaffungskosten der für die stereoskopischen Panoramenabbildungen notwendigen fünf Projektoren das vorgegebene Budget nicht übersteigen.

Nach der Art der Projektion auf die Leinwand wird zwischen *Auf- und Rückprojektion* unterschieden. Während sich bei der Aufprojektion Betrachter und Projektor vor der Projektionswand befinden, liegt der Projektor bei der Rückprojektion hinter der Leinwand. Aus diesem Grund sollte der Reflexionsgrad bei der Aufprojektion möglichst hoch und bei der Rückprojektion möglichst gering sein [23, S. 480]. Für die Projektion in der SEACEN-CAVE kann nur eine Aufprojektion zur Anwendung kommen, da die Rückprojektion einen komplexen, störanfälligen und durch den Einsatz großer Spiegel aufwendigen Aufbau erfordern würde. Entsprechend Abbildung 43 werden Projektionssysteme des Weiteren nach der Quelle der Lichterzeugung unterschieden. Während das Licht bei *aktiven Projektionssystemen* von einer durch das Videosignal gesteuerten Quelle ausgeht, wird das Licht bei *passiven Projektionssystemen* durch eine externe Lichtquelle erzeugt [19, S. 599].

Im Folgenden soll untersucht werden, welche aktiven bzw. passiven Projektionssysteme für die SEACEN-CAVE prinzipiell zum Einsatz kommen könnten. Dabei ist insbesondere die 3D-Fähigkeit der Projektoren zu berücksichtigen.

3.2.2.1 Aktive Projektionssysteme

Das Prinzip der zu den aktiven Systemen gehörenden *CRT-Projektoren*¹ beruht auf der Bildwiedergabe mittels Kathodenstrahlröhren. Dabei werden Leuchtstoffe zur Lichtemission durch das Auftreffen hoch beschleunigter Elektronen angeregt. Über eine optische Anordnung wird das so erzeugte Bild auf die Leinwand projiziert [23, S. 451, 481].

Drei kleine, monochromatische Kathodenstrahlröhren (rot, grün, blau) mit jeweils einem Strahlerzeuger werden dabei von ihrem farblich zugehörigen Signal angesteuert. Durch die Überlagerung der drei Primärfarbbilder entsteht durch additive Farbmischung ein farbiges Bild auf der Leinwand [19, S. 600] (Abbildung 44). Obgleich CRT-Projektoren Vorteile wie relativ hohe Helligkeit in dunklen Räumen, hoher Kontrast, HDTV- und 3D-Fähigkeit besitzen, überwiegen die durch die Röhren-Technologie bedingten Nachteile wie aufwendige Kalibrierung, geringe Transportabilität und hohe Anschaffungskosten [23, S. 481].

Ein weiteres aktives System wird durch *Laser-Projektoren* realisiert. Zum Einsatz kommen dabei stark gebündelte Lichtstrahlen aus drei für die Grundfarben benötigten Lasern, die - gesteuert durch ein Videosignal- additiv auf der Projektionsfläche abgebildet werden. [23, S. 482] (Abbildung 45). Trotz einiger Vorteile der Lasertechnologie, die kontrastreiche, hoch auflösende und scharfe Abbildungen ermöglicht, stehen Laser-Projektoren wegen der hohen Anschaffungskosten und der notwendigen Bedienung durch Fachpersonal für die SEACEN-CAVE nicht zur Wahl.

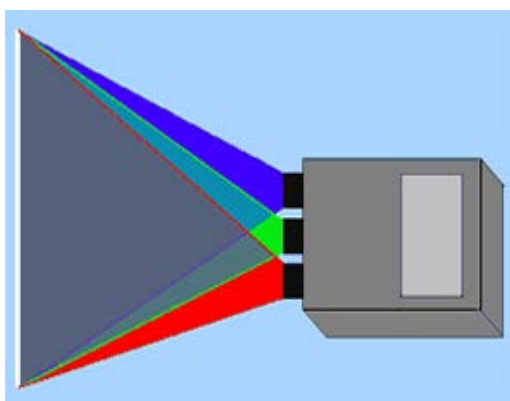


Abbildung 44: : Prinzip eines CRT-Projektors [47, Abb.]

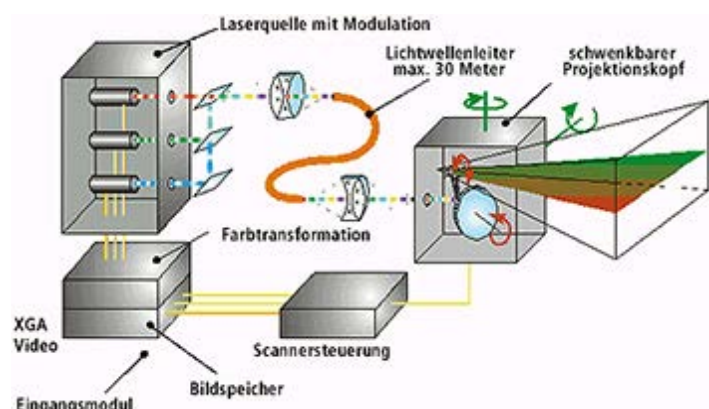


Abbildung 45: Prinzip eines Laser-Projektors [48, Abb.]

¹ CRT - Cathode Ray Tube; Kathodenstrahlröhre

Die Funktionsweise von *LED-Projektoren*¹ basiert auf der DLP-Technologie (Vgl. Abschnitt 3.2.2.2, S. 38). Im Gegensatz zu DLP-Projektoren² arbeiten LED-Projektoren jedoch als aktive Systeme, d.h. die Lichtquelle (LED) wird direkt vom Videosignal gesteuert. Während bei herkömmlichen DLP-Projektoren für die Farberzeugung ein Farbrad verwendet wird, kommen bei LED-Projektoren farbige LEDs der drei Grundfarben rot, grün und blau (RGB) als Lichtquelle zum Einsatz. Durch die Mischung der drei Grundfarben ist eine optimale Farbdarstellung möglich. Trotz vieler Vorteile, wie die Langlebigkeit der LEDs, der geringe Energieverbrauch, eine geringe Wärmeentwicklung und eine kompakte Bauweise ist die Lichtleistung von LED-Projektoren derzeit noch unzureichend. Aus diesem Grund kann diese Technologie nicht für das SEACEN-Projekt zum Einsatz kommen [49].

3.2.2.2 Passive Projektionssysteme

*DLP-Projektoren*² arbeiten wie die im letzten Abschnitt dargestellten LED-Projektoren mit digital gesteuerten, mechanischen Mikrosiegeln (DMD³). Da bei DLP-Projektoren eine externe Projektionslampe vom Videosignal gesteuert wird, sind diese Projektoren im Gegensatz zu LED-Projektoren den passiven Projektionssystemen zugeordnet. Das Licht der Projektionslampe, das durch ein rotierendes Farbrad hindurch auf den mit Mikrosiegeln versehenen Chip auftrifft, wird entsprechend der Kippwinkel jedes einzelnen Mikrosiegels in den Projektionsstrahlengang und dann auf die Leinwand oder in ein Absorptionsgebiet projiziert [23, S. 488-490]. Dabei wird jeder Mikrospiegel entsprechend der anliegenden Spannung, die einem bestimmten Grauwert entspricht, mehrere tausend Mal pro Sekunde ein oder ausgeschaltet (Abbildung 46). Mit Hilfe dieser Mikrospiegel-Technologie kann ein detailreiches Graustufenbild mit einer Abstufung von 1024 Grautönen erzeugt werden. Durch den Einsatz des rotierenden Farbrades kann ein Ein-Chip-DLP-System 16,7 Millionen Farben und ein Drei-Chip-System 35 Billionen Farben darstellen [50]. Da DLP-Systeme hohe Bildauflösungen und hohe Kontraste liefern, kommen sie insbesondere für hochwertige Geräte zum Einsatz [23, S. 489-490]. Da durch die sehr kurze Reaktionszeit von 16 Mikrosekunden das Nachziehen des Bildes vermieden wird, sind DLP-Projektoren für bewegungsreiche Szenen und dadurch auch für 3D-Projektionen geeignet [51].

Durch die Trennung der Lichterzeugung von der Adressierung der Bildelemente wird bei *LCD-Projektoren*⁴ als passiven Systemen die Projektion eines hellen und zugleich großen Bildes ermöglicht. Das bedeutet, dass die Leuchtdichte und die Auflösung unabhängig voneinander erhöht werden können [19, S. 602]. Bei der LCD-Projektion werden zunächst

¹ LED - Light Emitting Diode; Licht emittierende Diode

² DLP - Digital Light Processing; Projektionstechnologie der Firma Texas Instruments

³ DMD - Digital Micromirror Device; digitales Mikrospiegel Modul

⁴ LCD - Liquid Crystal Display; Flüssigkristallanzeige

Flüssigkristallanzeigen mit einer Hochleistungs-Metaldampflampe durchleuchtet. Dabei wird das Licht mit Hilfe von dichroitischen¹ Spiegeln in die drei Grundfarben rot, grün und blau zerlegt. Die Bündelung der drei Farbanteile erfolgt nach dem Durchdringen der Flüssigkristallanzeigen. Nach dem Passieren eines Linsensystems wird das gebündelte Licht mittels Projektionsobjektiv auf der Leinwand abgebildet (Abbildung 47). Neben hoher Lichtausbeute und hohen Kontrastwerten liegen die Vorteile von LCD-Projektoren vor allem in ihrer kompakten Bauweise. Des Weiteren erfolgt die Abbildung der drei Farbanteile -im Gegensatz zu Kathodenstrahl-Projektoren- über ein gemeinsames Linsen- bzw. Projektionssystem. Nachteilig hingegen ist die hohe Absorption der Lichtenergie im Flüssigkristallelement. Die Lichtleistung kann nicht beliebig erhöht werden, da durch die Erwärmung das Lichtventil zerstört werden kann. Ein weiterer Nachteil ist die Sichtbarkeit des Punktmusters der zur Bildpunktsteuerung benutzten aktiven Matrix auf der Projektionsfläche. Diese Nachteile können mittels einer neueren Technologie zum größten Teil beseitigt werden. Bei diesem weiterentwickelten System – dem so genannten Image Light Amplifier (ILA)- werden die LCD Elemente nicht durchstrahlt, sondern arbeiten mit Reflexion [23, S. 484-485].

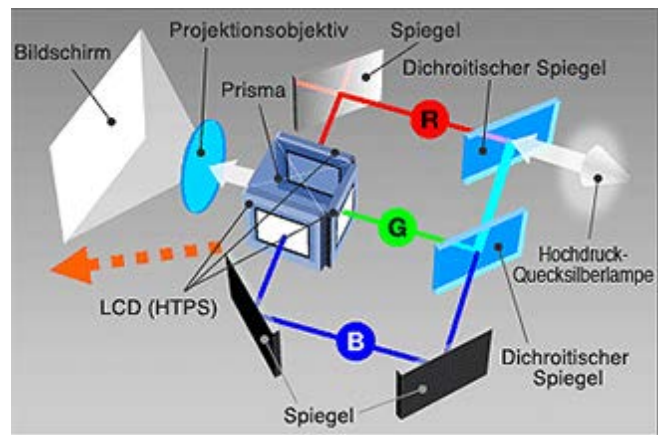
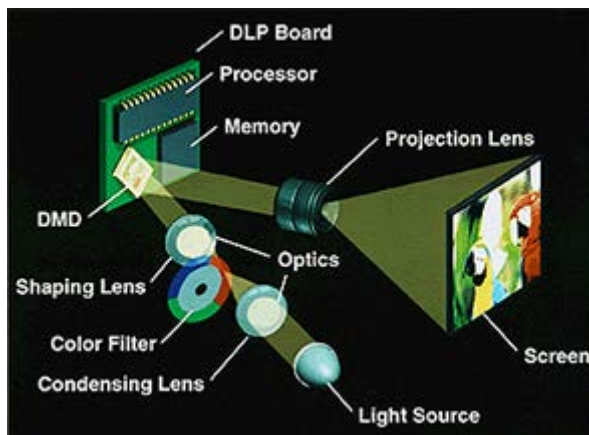


Abbildung 46: Prinzip eines DLP-Projektors [52, Abb.] **Abbildung 47:** Prinzip eines LCD-Projektors [53, Vgl. Abb.]

Ein ebenfalls auf direkt mit der Ansterelektronik verbundenen Flüssigkristallen basierendes passives System wird durch *LCoS-Projektoren*² realisiert. Diese Technologie funktioniert ähnlich wie bei ILA-Projektoren mittels Reflexion. Nach dem Passieren einer Glasplatte mit innen liegenden Elektroden und der sehr dünnen Flüssigkristallschicht trifft das Licht auf eine Vielzahl von spiegelnden Aluminiumgegenelektroden über die die Spannungen an die Bildpunkte gelangen. Die Helligkeit lässt sich durch die Variation der anliegenden Spannung regeln. Wenn keine Spannung anliegt, wird das Licht in die Lampe zurückgeworfen. Aus der direkten Verbindung mit einem Chip aus einkristallinem Silizium resultiert eine hohe Elektronenbeweglichkeit, die wiederum eine hohe Schaltgeschwindigkeit ermöglicht. Der Nachteil von LCoS-Projektoren ist die Sichtbarkeit einer feinen Pixelstruktur [23, S. 487].

¹ Spiegel mit frequenzselektivem Reflexionsverhalten

² LCoS - Liquid Crystal on Silicon, Flüssigkristall auf Silikonschicht

3.2.3 Verfahren der Stereoskopie

Die im vorangegangenen Abschnitt erläuterten optischen Aufnahme- und Wiedergabesysteme bilden die Grundlage für die im Folgenden dargestellten stereoskopischen Aufnahme- und Wiedergabeverfahren, die für die Erzeugung einer plausiblen virtuellen Umgebung unumgänglich sind. Das wesentliche Auswahlkriterium zur Nutzung dieser Verfahren für die SEACEN-CAVE war die Erzeugung höchster Plausibilität unter Einhaltung des Projektbudgets.

Die meisten Verfahren beziehen sich dabei auf die Nutzung der Querdissipation, die die wichtigste Komponente in der Wahrnehmung räumlicher Tiefe darstellt (Vgl. Abschnitt 2.2, S. 9-10). Aus der Querdissipation resultieren zwei stereoskopische Netzhaut-Teilbilder mit leicht unterschiedlicher Perspektive. Stereoskopische Systeme müssen demzufolge die Trennung und spätere Zuordnung des jeweiligen stereoskopischen Teilbildes zum zugehörigen Auge ermöglichen [1, S. 86].

Daraus ergibt sich der Einsatz einer Technologie zur stereoskopischen Aufnahme, die das Grundprinzip des menschlichen Sehapparates mit Hilfe von zwei Kameras in Kombination mit zwei Objektiven adaptiert. Für die Wiedergabe wäre die Nutzung eines Volumendisplays optimal, da diese auch in die Tiefe abbilden. Die Betrachtung wäre dann ohne die heute noch notwendige 3D-Brille möglich. Obgleich erste Ansätze mittels Laser bzw. der Holografie existieren, übersteigt gegenwärtig der Rechenaufwand die Umsetzung solcher Systeme [23, S. 498]. Die verschiedenen Verfahren sind in der folgenden Übersicht zusammengefasst:

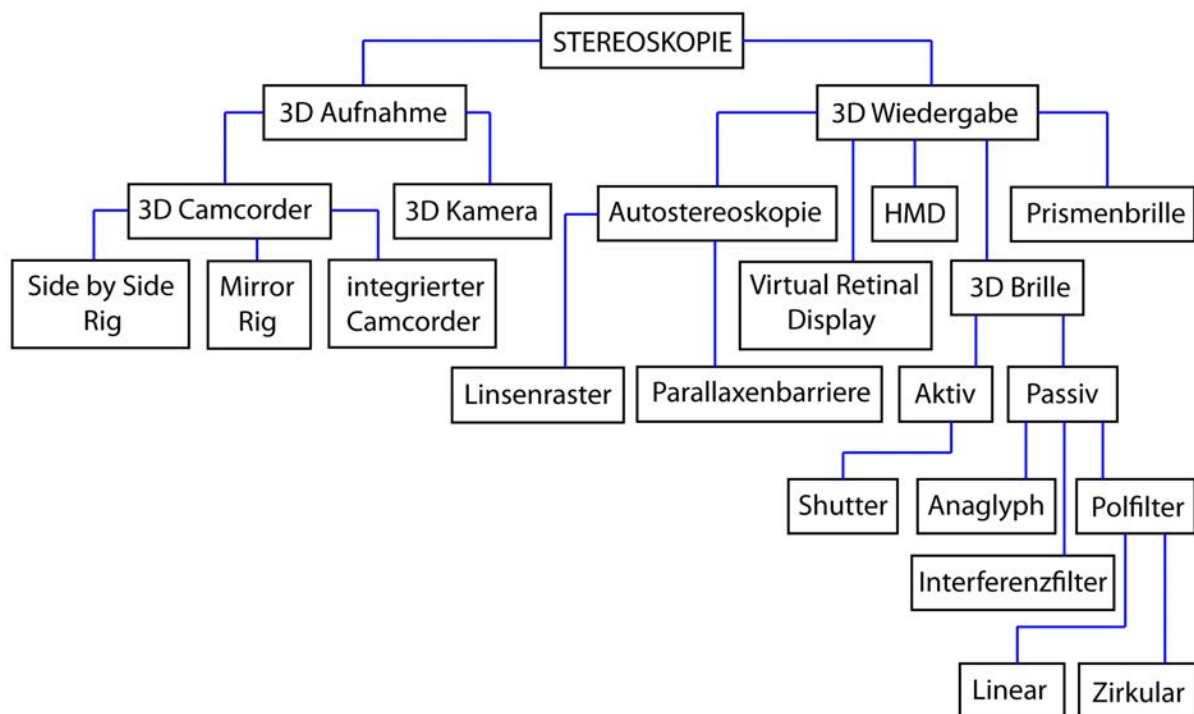


Abbildung 48: stereoskopische Aufnahme- und Wiedergabeverfahren

3.2.3.1 Verfahren für stereoskopische Filmaufnahmen

Die Aufnahme stereoskopischer Teilbilder kann auf einfache Weise mit Hilfe *einer* Kamera erfolgen. Dazu wird ein statisches Objekt nacheinander von zwei Positionen aus im ungefähren Augenabstand durch Verschiebung der Kamera senkrecht zur optischen Achse aufgenommen. Diese in den Anfangsjahren der Fotografie übliche stereoskopische Bildaufnahmetechnik [54, S. 21] kam ebenfalls für die Aufnahme der sechs Konzertsaal-Panoramen des SEACEN-Projektes mittels moderner Schiebe-/ Nodalpunktadapter zur Anwendung.

Um stereoskopische Filmaufnahmen zu realisieren, können zwei Camcorder bzw. ein Camcorder mit zwei Objektiven verwendet werden. Voraussetzung dafür ist, dass die optischen Achsen der Camcorder bzw. der Objektive im Augenabstand von ca. 65 mm voneinander entfernt liegen und die Aufzeichnungssysteme exakt gleich eingestellt und synchronisiert sind. Diese so genannten 3D-Rigs müssen darüber hinaus eine exakte Positionierung der Kameras gewährleisten. Während kleine Objektive bzw. Kameras mittels kleiner Gestelle auf einen Abstand (Stereobasis) zwischen 1 cm und 7 cm zusammengebracht werden können (Side by Side-Rig), werden stereoskopische Aufnahmen mit großen Kameras durch so genannte Mirror-Rigs realisiert, bei denen die beiden stereoskopischen Teilbilder über halbdurchlässige Spiegel zu den Kameras gelangen (Abbildung 49). Bei Mirror-Rigs kann die Stereobasis frei gewählt werden. Nachteilig hingegen ist ein durch die Spiegelung bedingter Lichtverlust von mindestens einer Blende [23, S. 448-449].

Die Integration von zwei Objektiven zu einem kompakten 3D-Camcorder führte seit 2010 zur Produktion verschiedener Modelle sowohl für den Konsumer- als auch für den professionellen Bereich.

Im Konsumer-Bereich gibt es Modelle mit nah nebeneinander liegenden Objektiven und echtem 3D Zoom (JVC GS-TD1, Sony HDR-TD10), aber auch Modelle die eine 3D Vorsatzlinse verwenden (Panasonic HDC-SD909 & VW-CLT1E). Die mittels dieser Technik erreichbare Bildqualität ist nicht befriedigend, da nur eine verringerte horizontale Auflösung zur Verfügung steht [55].

Im professionellen Bereich kommen seit 2010 kompakte, voll integrierte Full HD 3D-Camcorder zur Anwendung (Abbildung 50). Die zwei Objektive und das gesamte Aufnahmesystem mit zwei Speicherkarten sind dabei in einem leichtgewichtigen Gehäuse zusammengefasst. Während sich bei herkömmlichen Systemen stereoskopische Anpassungen nur mit Hilfe eines PCs oder externen Videoprozessors einstellen lassen, wird bei kompakten 3D-Camcordern durch stereoskopische Steuerungselemente eine einfache Handhabung und Bedienung ermöglicht, sodass u. a. der Konvergenzpunkt justiert werden kann. Die Speicherung der Videodaten erfolgt auf SDHC/SD Karten [56].

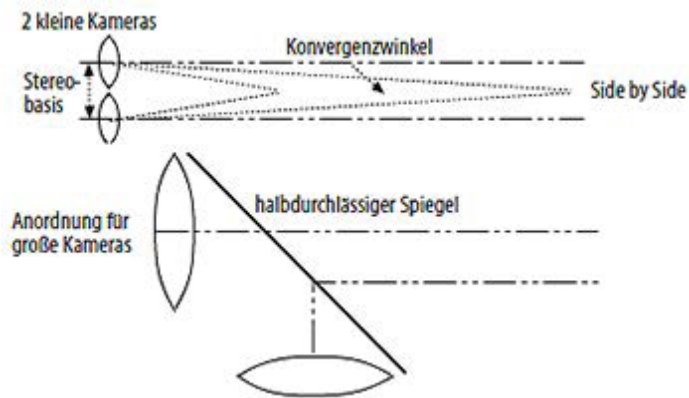


Abbildung 49: optische Anordnungen von Side by Side- und Mirror-Rigs [23, Abb. S.448]



Abbildung 50: Panasonic AG-3DA13D 3D-Camcorder mit integriertem Doppelobjektiv [57]

3.2.3.2 Verfahren für stereoskopische Filmwiedergabe

Um die separat aufgenommenen und gespeicherten stereoskopischen Teilbilder wiederzugeben, müssen diese dem jeweils zugehörigen Auge wieder zugeordnet werden.

Das erste aus einem Spiegelsystem bestehende Betrachtungsgerät für stereoskopische Bilder stellte Sir Charles Wheatstone 1838 der Öffentlichkeit vor [54, S. 19].

Außer der schon in Abschnitt 3.2.3 erwähnten, aber noch nicht realisierten Wiedergabe mittels eines Volumendisplay kommen gegenwärtig die noch mit Einschränkungen behafteten autostereoskopischen Systeme und die derzeit am weitesten verbreiteten Systeme mit 3D-Brille bzw. 3D-Helm zum Einsatz, die Im Folgenden vorgestellt werden.

Zu den ersten Systemen, die eine virtuelle Realität erzeugen konnten, gehören *Head Mounted Displays* (HMD). Bei diesen helmähnlichen Konstruktionen befinden sich direkt vor den Augen zwei kleine Displays, die das dem jeweiligen Auge zugeordnete Bild wiedergeben. Trotz der hohen Immersion durch die visuelle und akustische Abkopplung der Wahrnehmung des Benutzers von der Außenwelt, kommen derzeit überwiegend höher auflösende und ermüdungsfreier zu betrachtende Projektionssysteme zum Einsatz [1, S. 87]

Ein einfaches und wegen der geringen Anforderungen an die Hardware leicht zu bedienendes System mit 3D-Brille ist das *Anaglyphenverfahren*. Durch eine unterschiedliche Einfärbung der beiden stereoskopischen Teilbilder (meist rot und grün bzw. cyan) wird eine komplementäre Farbtrennung erzeugt. Nach Überlagerung der beiden Teilbilder kann jedes Auge mit Hilfe von Farbfilter-Brillen das ihm zugeordnete Teilbild betrachten (Abbildung 51). Nachteile dieses Verfahrens sind die aufgrund der Entfärbung entstandene eingeschränkte Farbdarstellung sowie das Auftreten von Geisterbildern durch die nur unvollständige Bildtrennung. Vorteilhaft hingegen sind die preiswerte und einfache Technik. Zur Anwendung kommt das Verfahren im Printbereich, bei Displays und bei Projektionen wie z.B. im Kino [23, S. 500].

Beim *Polarisationsverfahren* erfolgt durch die Nutzung der Polarisierungseigenschaften des Lichtes eine effektivere Bildtrennung als beim Anaglyphenverfahren. Trifft polarisiertes Licht auf einen senkrecht zur polarisierten Lichtebene stehenden Polarisator, kann das Licht nicht durch diesen hindurch gelangen. Die Trennung der stereoskopischen Teilbilder wird mit Hilfe von Polarisationsfolien realisiert. Diese Filter sind als zwei „Gläser“ in den Brillen um 90° gegeneinander gedreht angeordnet, sodass das Licht nur jeweils einer Polarisationsrichtung den jeweiligen Filter passieren kann. Alle anderen Polarisationsrichtungen werden nicht hindurch gelassen. Damit ist die gleichzeitige Projektion beider senkrecht polarisierter übereinander liegender Stereo-Teilbilder auf einer Leinwand möglich (Abbildung 52). Um die Polarisationsrichtung bei der Reflexion zu erhalten, kommen teure, silberbeschichtete Leinwände zum Einsatz. Nachteilig bei dieser so genannten linearen Polarisation ist die verminderte Trennung der Teilbilder bzw. des 3D-Effektes beim Neigen des Kopfes. Durch den Einsatz von rechts/links-zirkularer statt linearer Polarisation kann dieser Effekt vermieden werden. Das Polarisationsverfahren bietet zahlreiche Vorteile. So bleibt die Farbinformation erhalten und die Brillen sind preiswert und können ohne Batterie betrieben werden. Ein großer Nachteil ist hingegen ein durch den Einsatz von Polarisationsfiltern bedingter starker Lichtverlust. Des Weiteren müssen beim Polarisationsverfahren teure, polarisationserhaltende Leinwände verwendet werden [23, S. 500-501].

Im Bereich von Displays kommt das Polarisationsverfahren seit 2011 zum Einsatz. Im Unterschied zur Projektionsdarstellung können die polarisierten Stereo-Teilbilder jedoch nicht übereinander sondern nur abwechselnd zeilenweise auf dem Display dargestellt werden, wodurch sich die Auflösung halbiert. Die ungeraden Zeilen bilden dabei das rechte und die geraden Zeilen das linke Stereo-Teilbild [58].



Abbildung 51: Anaglyphenbild

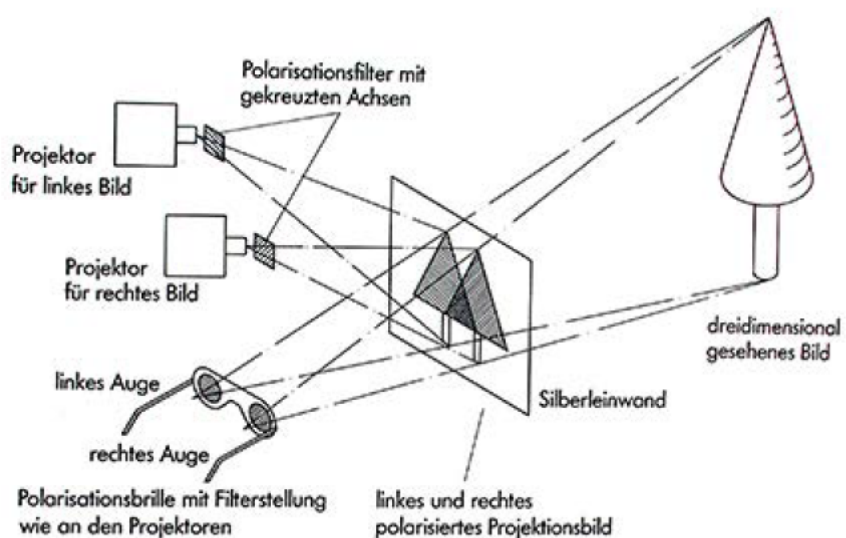


Abbildung 52: Polarisationsverfahren, Prinzip [59, Abb. S. 14]

Das *Shutterverfahren* arbeitet statt mit einer Farb- bzw. einer Polarisations- Bildtrennung mit einer zeitlichen und damit wirksameren Trennungsmethode der Stereo-Teilbilder. Bei diesem Verfahren werden spezielle Shutterbrillen verwendet, deren Flüssigkristallgläser elektronisch verschließbar sind. Die beiden Stereo-Teilbilder werden nacheinander auf der Leinwand bzw. dem Display abgebildet und gelangen durch das jeweils geöffnete Glas zu dem zugeordneten Auge, während das andere Glas jeweils elektronisch verschlossen wird. Damit jedes Stereo-Teilbild zum richtigen Zeitpunkt zu dem jeweils zugeordneten Auge gelangen kann, wird der Verschluss der Shutter-Brille z.B. durch Infrarotimpulse synchron zu den auf der Leinwand oder dem Display dargestellten Stereo-Teilbildern gesteuert [23, S. 501].

Da in einer Zeiteinheit nur eines der zwei Stereo-Teilbilder dargestellt wird, muss die Bildfrequenz verdoppelt werden. Dadurch wird die Wahrnehmung eines flimmerfreien Bildes ermöglicht (Abbildung 53). Moderne CRT-Monitore bzw. CRT-, LCD- oder DLP-Projektoren werden dabei mit Frequenzen von mindestens 120 Hz betrieben. Der Einsatz von CRT-Monitoren oder -Projektoren ist insbesondere durch das Nachleuchten der Displays problematisch, da dies zu einer verminderten Stereobildtrennung führen kann [1, S. 88].

Durch dieses Verfahren mit nur einem Projektor ohne spezielle Anforderungen kann eine gute Trennung der Stereo-Teilbilder und eine relativ preisgünstige, qualitativ hochwertige, stereoskopische Wiedergabe realisiert werden. Nachteilig sind dagegen die aufwändige 120 Hz Technik, die teure, energieabhängige Brillentechnik und der durch die dunklen Flüssigkristallgläser bedingte Lichtverlust [23, S. 29, 501].

Zur Anwendung des Shutterverfahrens bei Monitoren mit Zeilensprungverfahren kann die Umschaltfrequenz der Shutterbrille mit der Frequenz der Interlaced-Halbbilder synchronisiert werden. Diese nur aus den ungeraden oder nur aus den geraden Zeilen bestehenden Bilder müssen dabei den Stereo-Teilbildern entsprechen. Da jeweils nur die Hälfte der Zeilen für ein Bild zur Verfügung steht, halbiert sich die vertikale Auflösung [23, S. 501].

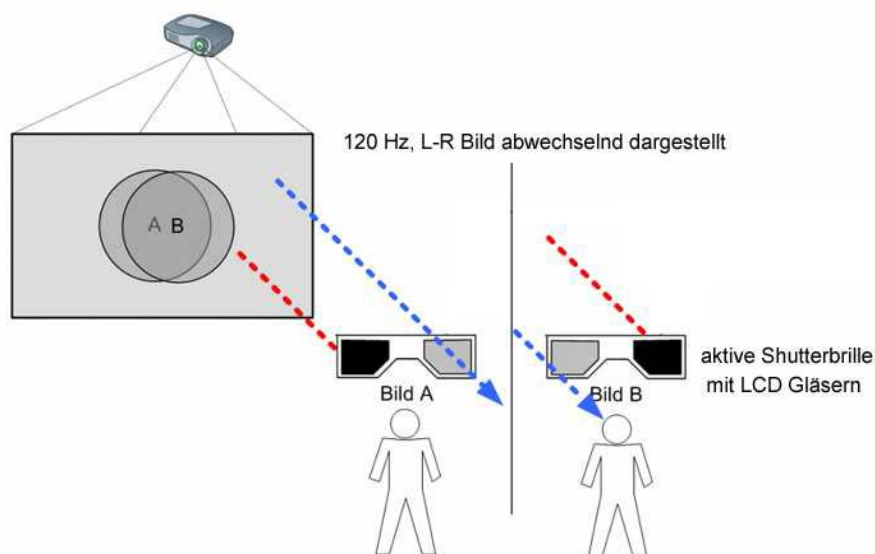


Abbildung 53: Shutterverfahren, Prinzip [60]

In abgewandelter Form kommt das Shutterverfahren für HDTV-fähige Displays und Projektoren, die auf der DLP-Technologie von Texas Instruments basieren, zum Einsatz. Aufgrund der schnellen Beweglichkeit der digitalen Mikrospiegel kann dabei die erforderliche, hohe Shutterfrequenz realisiert werden. Durch die schachbrettartige Trennung der beiden Stereo-Teilbilder kann ein flimmerfreieres Bild erzeugt werden, bei dem jedoch die Auflösung halbiert ist [61, S. 2-3].

Der entscheidende Vorteil der Shuttertechnologie bleibt die hohe Bildqualität, wodurch jedem Auge ein 1920 x 1080 Pixel großes Vollbild im so genannten Full-HD-3D Format bereitgestellt werden kann [62].

Neben diesen populären stereoskopischen Wiedergabeverfahren halten u.a. im Kino-Bereich neue Entwicklungen wie das *Wellenlängenmultiplex-Verfahren* Einzug. Dieses von Infitec entwickelte und durch die Dolby 3D Lizenzierung zunehmend verbreitete Verfahren basiert auf einer Aufspaltung der einzelnen RGB Frequenzanteile [63]. Zur Trennung der Anteile werden abweichende Grundfarben verwendet, die jedoch denselben Farbeindruck erzeugen können. Dabei kommt für jedes Stereo-Teilbild jeweils ein mit schmalbandigen Interferenzfiltern ausgestatteter Projektor zum Einsatz, der nur die jeweiligen leicht unterschiedlichen Grundfarben eines sehr schmalen, aber unterschiedlichen Frequenzspektrums passieren lässt. Mittels der in den Brillengläsern entsprechenden schmalbandigen Interferenzfilter wird die Trennung der beiden auf der Leinwand übereinander projizierten Bilder ermöglicht (Abbildung 54) [1, S. 89]. Neben dem batteriefreien Betrieb der Brillen bietet das Verfahren laut Hersteller Infitec zahlreiche Vorteile, wie die uneingeschränkte Farbdarstellung, die Kompatibilität mit Standard-Kinoleinwänden und weißen Wänden jeglicher Art sowie die Kompatibilität mit den gängigen Projektortechnologien (u.a. LCD, DLP) [64, S.8].

Ein System, das auf den ersten Blick Ähnlichkeiten zur HMD-Technik aufweist, ist das *Virtual Retinal Display*. Es wird ebenfalls wie eine Brille vor das Auge platziert, arbeitet jedoch nach einem anderen Prinzip. Mit Hilfe eines modulierten Laserstrahls wird das Bild im Gegensatz zu allen bisherigen Verfahren direkt auf die Netzhaut projiziert. Nach Analyse des Eingangssignals über eine Steuerelektronik wird durch Helligkeitsmodulation des Laserlichts von einem horizontalen und einem vertikalen Scanner ein Laserbild erzeugt. Da dieses Bild noch zu klein ist, wird es durch eine Linse aufgefächert und gelangt über einen halbdurchlässigen Spiegel auf die Netzhaut. Durch die dadurch entstehende Abdunklung des realen Bildes wird eine Integration von virtuellem und realem Bild erreicht (Abbildung 55). Die Vorteile dieses Systems liegen in der direkten Bilderzeugung auf der Netzhaut, die eine hohe Auflösung und Helligkeit ermöglicht. Dadurch kann das System auch in heller Umgebung zur Anwendung kommen. Des Weiteren wird der bekannte Konflikt zwischen Akkomodation und Konvergenz durch die Entstehung eines relativ natürlichen 3D Eindruckes verringert. Neben dem Potential als Sehhilfe gibt es zahlreiche Einsatzmöglichkeiten von Virtual Retinal Displays z.B. in der Industrie, Medizin, Kommunikation und in den Medien [11, S. 4-11].

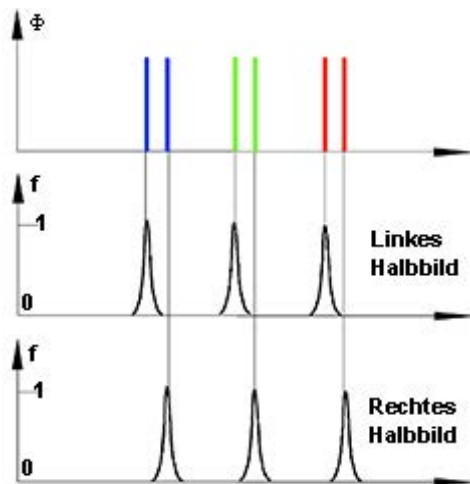


Abbildung 54: Wellenlängenmultiplex Verfahren, Prinzip [1, Abb. S. 89].

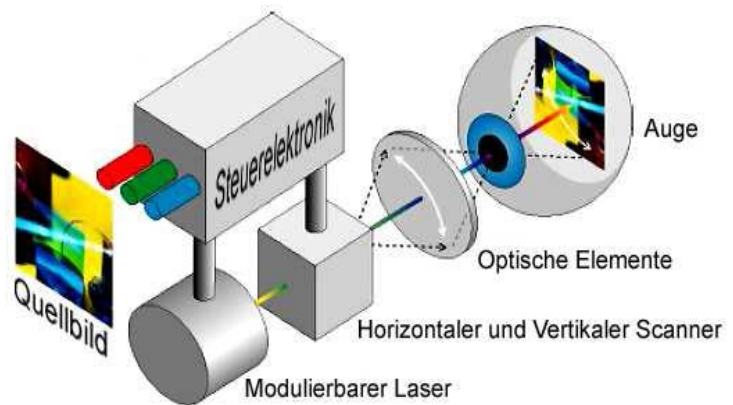


Abbildung 55: Virtual Retinal Displays – VRD¹, Prinzip [11, Vgl. Abb. S. 4]

Ein weiteres Verfahren das nur mittels Apparatur vor den Augen funktioniert, ist das *KMQ-Verfahren*.² Diese aus den Anfängen der Stereoskopie stammende Technik basiert auf dem Einsatz von Prismen als Strahlableitung. Die räumlich übereinander oder nebeneinander getrennt dargestellten beiden Stereo-Teilbilder werden mit Hilfe einer Prismenbrille zu einem stereoskopischen Bild fusioniert. Wegen der Nachteile wie halbierte Auflösung, feste Betrachtungsposition und –abstand kommt dieses Verfahren nur bei einigen Spezialanwendungen zum Einsatz [65, S.88-89].

Entsprechend ihren Eigenschaften wird bei den zuvor beschriebenen Systemen zwischen aktiven und passiven stereoskopischen Verfahren unterschieden. Während bei den passiven Verfahren wie Anaglyphen-, Polarisations- und Wellenlängenmultiplexverfahren die Stereo-Teilbilder permanent dargestellt und mit Hilfe optischer Filter ohne Energieversorgung getrennt werden, kommen bei aktiven Systemen, wie Shutterverfahren oder Virtual Retinal Display, elektronisch gesteuerte Wiedergabegeräte mit einer Spannungsversorgung und aktiven, batteriebetriebenen Brillen zum Einsatz. Diese wegen der elektronischen Brillen und der hohe Bildwiederholfrequenz der Projektoren bzw. Displays teuren aktiven Systeme bieten in der Regel eine bessere Bildqualität und Stereobildtrennung als die passiven Systeme [5, S.22-23].

Im Gegensatz zu den bisher dargestellten Wiedergabeverfahren ist die *autostereoskopische Wiedergabe* nicht an Brillen oder andere Sehhilfen gebunden. Die Trennung der Stereo-Teilbilder erfolgt bei diesem in zwei Varianten existierenden Verfahren direkt am Bildschirm. Das Display besteht aus schmalen vertikalen Streifen, die abwechselnd dem jeweils zugeordneten Auge dargeboten werden. Eine Umsetzung dieser Technologie wird mittels einer Parallaxenbarriere realisiert. Dabei gelangen von einem bestimmten geeigneten

¹ VRD - Virtual Retinal Display; lasergestütztes Abbildungssystem auf der Netzhaut

² abgeleitet von den Initialen der Erfinder Koschnitzke, Mehnert und Quick

Betrachtungsort aus nur die für das passende Auge bestimmten Stereo-Teilbild-Streifen zu dem jeweils passenden Auge. Die jeweils nicht passenden Stereo-Teilbild-Streifen werden mit Hilfe der Parallaxenbarriere ausgeblendet (Abbildung 56). Eine andere Technik der Bildtrennung funktioniert mittels Linsenraster. Durch die geeignete Krümmung der vertikal verlaufenden Linsen können von einem festen Betrachtungsort aus die Stereo-Teilbild-Linsenstreifen zu dem jeweils passenden Auge gelenkt werden. Da bei diesen Verfahren das Gesamt-Stereobild aus zwei verkämmteten Stereo-Teilbildern besteht, halbiert sich die horizontale Auflösung. Ein weiterer Nachteil ist, dass die Bildbetrachtung nur von einem bestimmten Ort aus und damit nur für eine Person möglich ist. Mit Hilfe von Verfolger-Systemen (Tracking) die den aktuellen Betrachterstandort mittels Kameras registrieren und die Linsenstreifen-Winkel entsprechend nachregeln, soll zumindest die Betrachtung von verschiedenen Betrachtungsorten aus für eine Person ermöglicht werden. [23, S. 499-500].

Um eine Darstellung für mehrere Betrachter bzw. Blickwinkel zu ermöglichen, wird jedes Pixel in Subpixel geteilt: Die verschiedenen Subpixel können über ein Linsenraster bzw. eine Parallaxenbarriere zu verschiedenen Betrachtungsorten gelenkt werden, wodurch eine stereoskopische Betrachtung des Displays von mehreren Personen an verschiedenen Standorten aus möglich wird (Abbildung 57). Mit einer Zahl von ca. 100 Subpixeln kann eine ausreichende Stereoqualität erzeugt werden. Jedoch verringert sich mit jeder weiteren möglichen Blickrichtung die horizontale Auflösung. Bisher konnte eine Zahl von 7 Subpixeln realisiert werden. Für eine befriedigende autostereoskopische Darstellung ohne Sehhilfen ist demzufolge noch viel Entwicklungsarbeit notwendig [1, S. 90].

Bei der Kombination verschiedener Techniken wie beispielsweise Trackingsystemen mit der Erzeugung hoher Subpixelzahlen kommt in naher Zukunft möglicherweise eine akzeptable autostereoskopische Wiedergabe für eine begrenzte Zahl an Betrachtern (z.B. im familiären Bereich), die ihren Betrachtungswinkel ändern können, zum Einsatz.

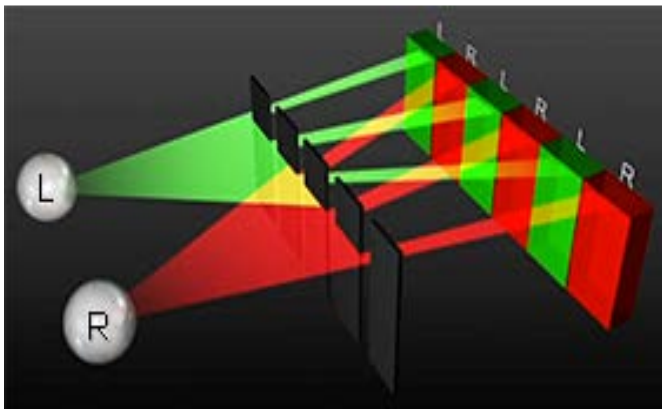


Abbildung 56: Autostereoskopie; Parallaxenbarriere [66, Abb. aus Video]

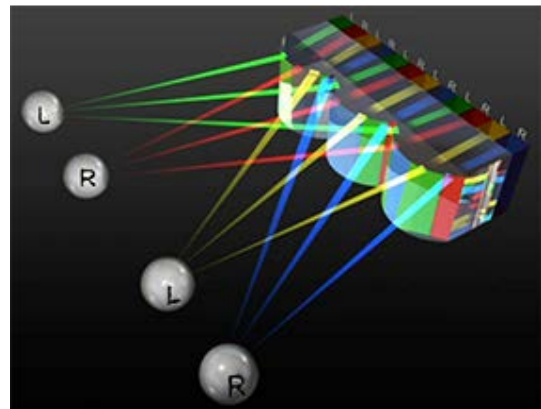


Abbildung 57: Autostereoskopie; Linsenraster für mehrere Betrachter [66, Abb. aus Video]

3.3 Grundlagen und Grenzen der auditiven Wahrnehmung

Um eine virtuelle Realität zu erzeugen, die eine plausible auditive Wahrnehmung ermöglicht, werden Kenntnisse über die Möglichkeiten und Grenzen des menschlichen auditiven Systems und dessen Eigenschaften benötigt. Die grundlegende Voraussetzung für die auditive Wahrnehmung ist die Übertragung von Schallereignissen zum Ohr.

3.3.1 Schall

Schall ist eine auditiv wahrnehmbare Druckänderung in der Luft oder einem anderen Medium. Wenn durch Bewegungen oder Vibrationen eines Objektes in dem umgebenden Medium Druckänderungen hervorgerufen werden, entsteht ein Schallereignis und damit ein Schallstimulus für ein auditives System [4, S. 259]. Die dabei erzeugten kleinen Schwingungen werden durch die Kompressibilität und die Masse der Luft bzw. des Mediums übertragen und gelangen schließlich zum Ohr. Die dabei entstehenden Druckschwankungen sind dem atmosphärischen Ruhedruck p_0 überlagert und charakterisieren die wichtigste zeit- und ortsabhängige akustische Feldgröße - den Schalldruck p . Ein Schallereignis wird durch die zwei wesentlichen Eigenschaften Klangfarbe und Lautstärke charakterisiert. Während die Lautstärke durch den Schalldruck repräsentiert wird, ist die Frequenz¹ f ein Maß für die Klangfarbe bzw. die Tonhöhe. Das auditive System ist in der Lage Frequenzen zwischen 16 und 16.000 Hz wahrzunehmen. Der genaue Bereich hängt jedoch von verschiedenen Faktoren wie Lebensalter oder Dauerbelastung durch Lärm bzw. laute Musik ab. Der angegebene Wert bezieht sich auf einen gesunden Menschen von 20 Jahren. Von da an verringert sich die obere Grenze um ca. 1 kHz pro Lebensdekade. Ähnlich der zeitlichen Auflösung beim visuellen System existiert beim auditiven System am unteren Frequenzbereich die so genannte *Flimmergrenze*. Diese individuell leicht unterschiedliche bei ca. 16 Hz liegende Grenze ist ein Maß dafür, ab wann eine Ereignisfolge als zusammenhängendes Schallereignis wahrgenommen wird. Insbesondere setzen sich natürliche Schallereignisse ähnlich wie das Licht aus verschiedenen Frequenzen des gesamten Spektrums zusammen. Umgekehrt sind beliebige Signale durch die Summe von reinen Tönen mit unterschiedlicher Frequenz bzw. Amplitude darstellbar (Abbildung 58). Daraus resultiert, dass akustische Signalübertragungen in der Regel mit Hilfe von Frequenzgängen beschrieben werden [20, S. 1-3]. Eine wichtige Eigenschaft des auditiven Systems ist die relative Schallwahrnehmung. Demnach werden Reizänderungen als gleich empfunden wenn diese im gleichen Verhältnis geändert werden. Aus diesem Zusammenhang resultiert das *Weber-Fechner-Gesetz*, das in folgender Gleichung zusammengefasst ist [20, S. 5, 15]:

¹ Anzahl der Schwingungen pro Sekunde; Einheit: Hertz (Hz)

$E = 2,3 k \lg \frac{R}{R_0}$	E = Empfindung R = Reiz R₀ = Reizschwelle k = Konstante
-------------------------------	---

Ähnlich wie bei der Helligkeitsempfindlichkeit des visuellen Systems (Vgl. Abschnitt 3.1.2, S. 19) funktioniert auch die auditive Wahrnehmung nach diesem logarithmischen Gesetz. Dadurch sind Schalldrücke eines ca. sieben Zehnerpotenzen umfassenden Bereiches erfahrbar. Während schwache Reize sehr gut wahrnehmbar sind, werden große Reize stark abgeschwächt wahrgenommen. Dabei wirkt die logarithmische Kennlinie wie ein „Überlastschutz“ und ermöglicht die insbesondere für das frühe Überleben des Menschen vorteilhafte Wahrnehmung eines großen Wertebereiches von Schalldrücken. Entsprechend diesen Erkenntnissen wird für das technische Maß ebenfalls eine logarithmierte Größe verwendet:

$L = 20 \lg \frac{p}{p_0}$	L = Schalldruckpegel p = Schalldruck p₀ = Hörschwelle
----------------------------	---

Dieser gut handhabbare Schalldruckpegel **L** bildet den erfahrbaren Wertebereich des physikalischen Schalldruckes auf einer Skala von etwa 0 bis 140 dB ab. Die Bezeichnung dB (Dezibel) weist auf die Verwendung des logarithmischen Bildungsgesetzes hin [20, S. 6-7].

Das zuvor beschriebene Weber Fechner Gesetz gilt auch für die Empfindung der Tonhöhe bzw. der Frequenz. So werden beispielsweise Oktavsprünge (12 Halbtonschritte) immer als gleichgroßer Frequenzsprung empfunden. Misst man jedoch die Frequenzen zwischen mehreren Oktaven, so erhöhen sich die Frequenzen nicht jeweils um einen konstanten Frequenzwert sondern verdoppeln sich bei jedem Oktavsprung [20, S. 4].

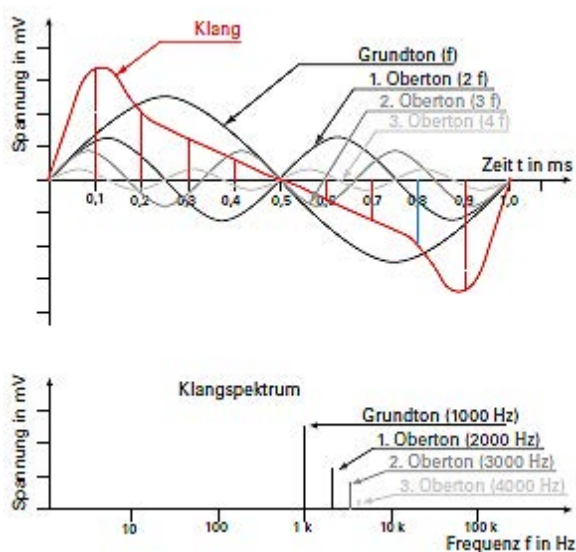


Abbildung 58: Zeit- und Frequenzverhalten eines einfachen Klanges [28, Abb. S. 955]

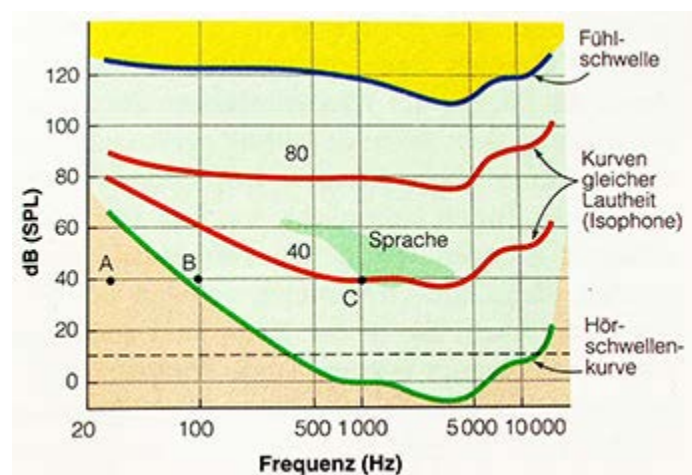


Abbildung 59: Hörfläche durch Hörschwellenkurve und Föhlschwelle begrenzt [4, Abb. S. 263]

Wie bereits beschrieben, ist die auditive Wahrnehmung des Menschen wie auch beim visuellen System auf einen schmalen Frequenzbereich zwischen 16 Hz und 16.000 Hz begrenzt – den so genannten *Hörbereich* [20, S. 1]. Die Empfindlichkeit des Gehörs ist jedoch frequenzabhängig. Am empfindlichsten ist das auditive System zwischen 2000 und 4000 Hz. Dieser Bereich ist insbesondere für die Sprachwahrnehmung bedeutsam. Daraus resultiert die so genannte *Hörfläche*, die den Bereich des menschlichen Hörens darstellt (Abbildung 59). Begrenzt wird diese von der unten befindlichen Hörschwellenkurve und der oben liegenden Föhlschwelle. Unterhalb der Hörschwelle sind keine Schallereignisse wahrnehmbar, oberhalb der Föhlschwelle werden hingegen Schmerzen empfunden. Aus der Frequenzabhängigkeit resultieren auch die so genannten *Isophone*. Diese Kurven gleicher Lautheit entstehen dadurch, dass unterschiedlich große Schalldrücke (hier in dB-Werten) bei verschiedenen Frequenzen als gleich laut empfunden werden [4, S. 262-263].

3.3.2 Physiologische Grundlagen

Die Erfassung, Umwandlung, Übertragung und Wahrnehmung von Schall werden durch das auditive System ermöglicht. Dabei wird der Schallstimulus zu den Rezeptoren geleitet, in elektrische Signale umgewandelt und so verarbeitet, dass Stimulus-Eigenschaften wie Lautheit, Tonhöhe, Klangfarbe und Richtung wahrgenommen werden können. Physiologisch betrachtet, wird zwischen dem äußeren Ohr, dem Mittel- und dem Innenohr unterschieden (Abbildung 60)[4, S. 267]. *Das äußere Ohr* leitet die Schallwellen über die Ohrmuschel und den äußeren Gehörgang, der eine Schutzfunktion und Frequenzverstärkung durch Resonanz gewährleistet, zum Trommelfell weiter. Dieses überträgt die Schwingungen an die im ca. zwei Kubikzentimeter großen *Mittelohr* befindlichen drei Gehörknöchelchen: Hammer, Amboss und Steigbügel. Von dort aus gelangen die Schwingungen über eine Membran und das ovale Fenster in das *Innenohr*, das im Wesentlichen aus der schneckenförmigen, flüssigkeitsgefüllten Cochlea besteht (Abbildung 61).

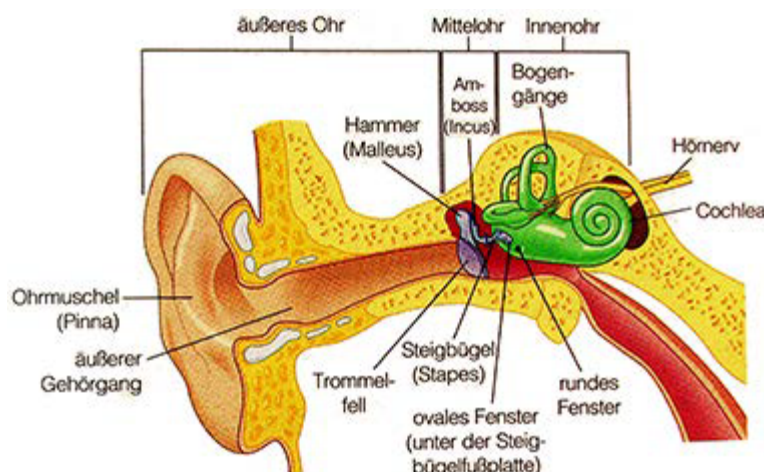


Abbildung 60: Das Ohr des Menschen [4, Abb. S. 267]

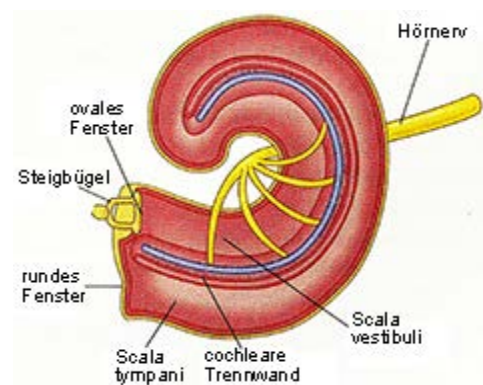


Abbildung 61: Cochlea [4, Abb. S. 269]

Wegen der höheren Dichte der cochlearen Flüssigkeit im Innenohr gegenüber der Luft im Mittelohr erfolgt mit Hilfe der Gehörknöchelchen eine Anpassung bzw. Verstärkung ähnlich dem Hebelprinzip, indem die Schwingungen vom großflächigen Trommelfell auf den viel kleineren Steigbügel übertragen werden. Die im Innenohr befindliche Cochlea ist durch die cochleare Trennwand, die das Corti'sche Organ enthält, in zwei längliche Abschnitte untergegliedert. Das Corti'sche Organ besteht aus der Basilarmembran, der Tektorialmembran und den äußeren und inneren Haarzellen, die die Rezeptoren für das Hören darstellen. Die vom Steigbügel auf die cochleare Flüssigkeit übertragenen Schwingungen bewirken eine Auf- und Abwärtsbewegung der cochlearen Trennwand. Dadurch wird das Corti'sche Organ in eine Relativbewegung versetzt (Abbildung 62, blauer Pfeil) während gleichzeitig eine Vorwärts- und Rückwärtsbewegung der Tektorialmembran (Abbildung 62, roter Pfeil) entsteht. Aufgrund dieser zwei Bewegungen werden die an den oberen Enden der inneren Haarzellen ansetzenden Stereozilien ausgelenkt, da sie sich in der sie umgebenden Flüssigkeit bewegen und gleichzeitig die Tektorialmembran berühren. Durch diese Auslenkung der Stereozilien werden elektrische Signale erzeugt, die an Nervenfasern im Hörnerv übertragen werden. Dabei genügen bereits Bewegungen von 100 Billionstel Meter um eine Antwort der Haarzelle und damit ein elektrisches Signal auszulösen [4, S. 267-271].

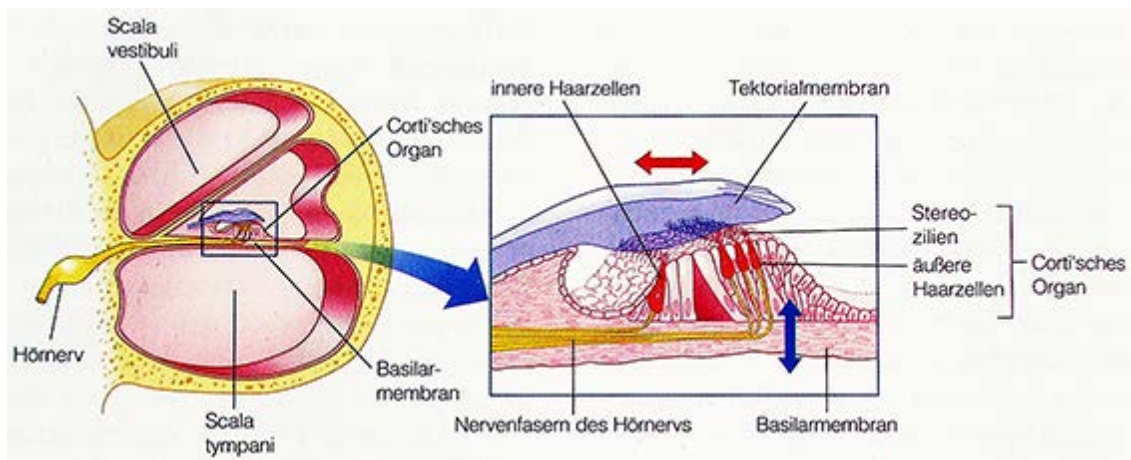


Abbildung 62: links: Querschnitt durch die Cochlea; rechts Corti'sches Organ (vergrößert) [4, Abb. S. 270]

3.3.3 Auditive Wahrnehmung

Nach einer Reihe von Studien fand der Physiologe Békésy heraus, dass die Cochlea von einer Wanderwelle durchlaufen wird und dass die Frequenz eines Tones offenbar mit einem örtlich bestimmten Abschnitt der Basilarmembran korreliert. Dieser als Ortstheorie bekannte Zusammenhang führte zu einer *tonotopen Karte* der Cochlea auf der die Frequenzen wie auf einer Skala regelmäßig angeordnet sind. Danach reagiert der Apex¹ auf tiefe Frequenzen dagegen die Basis in der Nähe des Steigbügels nur auf hohe Frequenzen (Abbildung 63).

¹ Spitze einer Schnecke; hier: Zentrum bzw. Ende der Cochlea

Die mit den Haarzellen verbundenen Neuronen antworten nur innerhalb eines sehr schmalen Frequenzbereiches woraus die so genannten *Frequenz-Tuningkurven* resultieren. Dementsprechend kann die Cochlea unter Einbeziehung der Frequenz-Tuningkurven als Gruppe von Filtern betrachtet werden. Trifft ein komplexer Klang auf die Basilarmembran, wird dieser in seine spektralen Komponenten zerlegt, d.h. jeder Abschnitt der Basilarmembran, der am meisten mit einer bestimmten Frequenz korreliert, reagiert mit einer neuronalen Antwort (Abbildung 64). Letztlich führt die Basilarmembran eine so genannte *Fourieranalyse* durch. Neben dieser örtlichen Frequenz-Analyse durch das auditive System gibt es auch einen zeitbezogenen Zusammenhang. Danach korrespondiert der zeitliche Verlauf der in Salven feuernden Nervenfasern mit der Frequenz des Schallstimulus. Dabei tritt die so genannte Phasenkopplung auf, die die Neuronen immer synchron zum maximalen Phasenwert des Schallsignals feuern lässt. Aus diesen Forschungserkenntnissen lässt sich schließen, dass sowohl eine örtliche als auch eine zeitliche Analyse der Frequenz im Wahrnehmungsprozess genutzt wird. Diese durch Nervenfasern umgewandelte Frequenzinformation gelangt über weitere neuronale Strukturen schließlich zum auditiven Kortex. Dieser besteht aus dem primären auditiven Kortex mit seinem Gürtel und entsprechend jüngeren Forschungsergebnissen zufolge einem sekundären auditiven Kortex (erweiterter Gürtel). Die Verarbeitung der Schallstimuli erfolgt zunächst im Kernbereich. Komplexe Klänge wie Rauschen oder Sprache aktivieren zusätzlich den Gürtel bzw. den erweiterten Gürtel. Neuere Forschungen belegen, dass weitere, zusätzliche Areale im Kortex auf auditive Stimuli reagieren. In Bezug zu Untersuchungen cross-modaler Überschneidungseffekte im Rahmen des SEACEN-Projektes ist die Existenz von bestimmten Arealen, die sowohl durch visuelle als auch durch auditive Stimuli aktiviert werden können von besonderem Interesse. Ein Teil dieser Areale ist über zwei voneinander getrennte Verarbeitungsströme miteinander verbunden – den *Was- und Wo-Strömen* [4, S. 272-281].

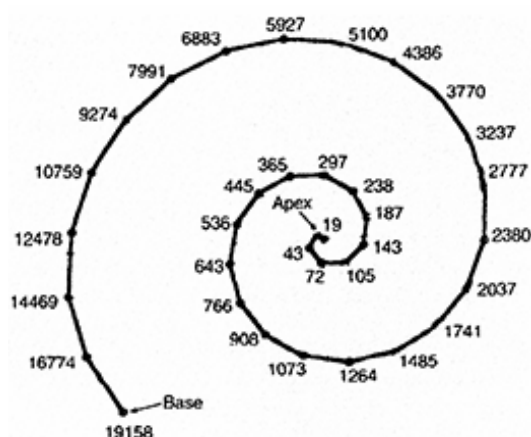


Abbildung 63: Tonotopie Karte der menschlichen Cochlea mit Frequenzzuordnung (in Hz) zu den Orten der max. neuronalen Antwort [67]

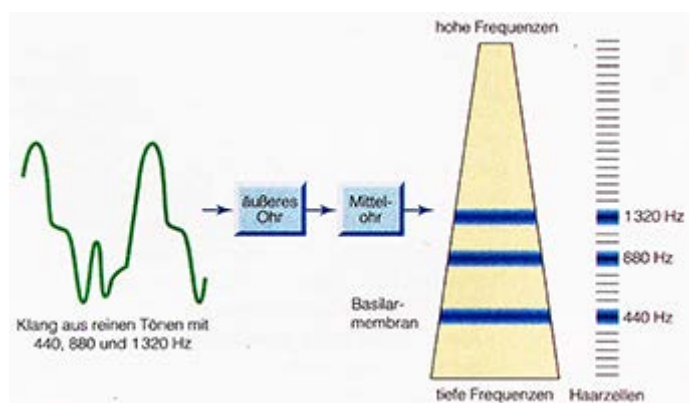


Abbildung 64: die Cochlea zerlegt Schallereignisse in Frequenzanteile (Frequenzanalysator) durch Stimulation unterschiedliche Areale der Basilarmembran [4, Abb. S. 278]

Während über den Was-Strom die Identifizierung eines Schallereignisses erfolgt, ist durch den Wo-Strom die Lokalisation der Schallquelle möglich. Für letzteres spielt insbesondere der primäre Kortex eine entscheidende Rolle wie auch für die Verarbeitung der Frequenz. So existiert offenbar ähnlich wie in der Cochlea auch im primären Kortex eine tonotope Karte, die bestimmte Frequenzen bestimmten kortikalen Arealen zuordnet [4, S. 282-283].

3.3.4 Auditive Lokalisation

Das menschliche auditive System ist in der Lage Schallquellen innerhalb des wahrnehmbaren, so genannten *auditiven Raumes* zu lokalisieren. Dieser auditive Raum kann mit Hilfe von drei Koordinaten vermessen werden. Während das *Azimuth* den Winkel in der horizontalen Ebene angibt, misst die *Elevation* den Winkel innerhalb der vertikalen Ebene. Die *Entfernung* gibt die Distanz des Zuhörers zur Schallquelle an. Während beim visuellen System Objekte, die sich an verschiedenen Orten befinden, auf unterschiedlichen Arealen der Netzhaut getrennt abgebildet werden, erfolgt die Trennung von unterschiedlich positionierten Schallquellen durch die Berechnung der *interauralen Zeit- und Pegeldifferenz*, die zwischen beide Ohren auftritt. Als Grundlage dieser auch als *binaurale Positionsreize* bezeichneten Stimuli dient der Vergleich der Schallsignale beider Ohren. Aus den dabei ermittelten Zeit- bzw. Schalldruckpegeldifferenzen kann die Position der Schallquelle berechnet werden (Abbildung 65). Die interaurale Pegeldifferenz tritt jedoch nur bei hohen Frequenzen auf. Des Weiteren ist sie vom Einfallswinkel des Schalls abhängig. Da der Kopf ein Hindernis für Frequenzen über ca. 1000 Hz darstellt, kommt es je nach Einfallswinkel zur Abschwächung bestimmter höherer Frequenzen (Abbildung 66) [4, S. 291-295].

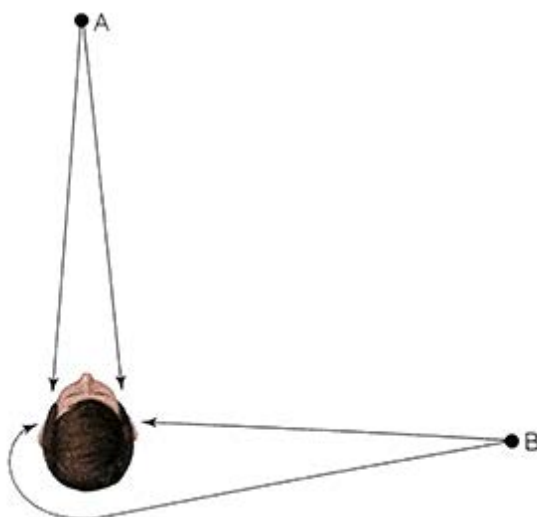


Abbildung 65: A: Schall erreicht beide Ohren gleichzeitig - keine Zeitdifferenz; B: Schall erreicht zuerst linkes Ohr - interaurale Zeitdifferenz tritt auf [4, Abb. S. 293]

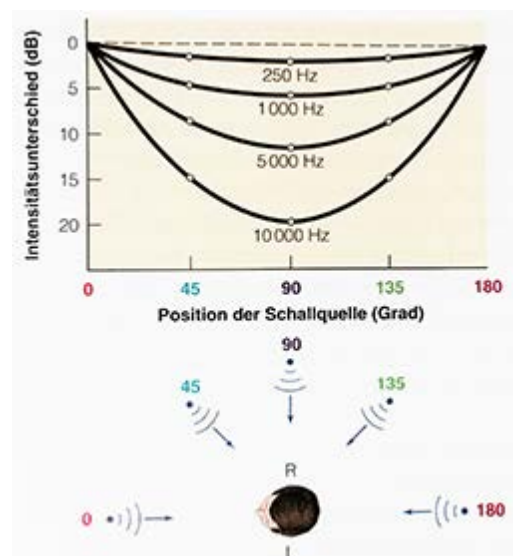


Abbildung 66: interaurale Pegeldifferenz unterschiedlicher Frequenzen bei verschiedenen Schalleinfallswinkeln [4, Abb. S. 295]

Anders als binaurale Positionsreize wirken *monaurale Positionsreize* bereits mit einem Ohr und nur in der vertikalen Ebene – bei unterschiedlichen Elevationen. Sie ergeben sich aus der spektralen Veränderung des Schalls durch Einfallswinkel-abhängige Reflexionen am Kopf und innerhalb der Ohrmuschel. Der tatsächlich in die Gehörgänge gelangende Schall wird durch die kopfbezogene Übertragungsfunktion, die so genannte *Head Related Transfer Funktion* (HRTF) beschrieben. Letztlich ergeben sich bei verschiedenen vertikalen Einfallswinkeln unterschiedliche Frequenzmuster, aus denen das auditive System auf die Position der Schallquelle schließen kann. Obgleich das auditive System alle drei Hinweisreize zur Lokalisation von Schallquellen nutzt, geht aus Experimenten von Wightman und Kistler (1997) hervor, dass die interaurale Zeitdifferenz bei hohen Frequenzen hingegen die interaurale Pegeldifferenz bei tiefen Frequenzen wirksam ist. Beide lokalisieren innerhalb der Azimutebene während die kopfbezogene Übertragungsfunktion monaurale, spektrale Positionshinweise in der Elevationsebene liefert. Der auditive Raum wird durch Antworten spezifischer Neuronen auf interaurale Zeitdifferenzen einerseits und durch die Existenz räumlicher, so genannter *topographischer Karten* im auditiven Kortex andererseits repräsentiert. Dabei korrespondieren bestimmte Orte in neuronalen Strukturen systematisch mit Positionen im Raum. Bei Untersuchungen im Kortex eines Affen konnte jedoch keine topographische Karte gefunden werden. Stattdessen wurden *Panoramaneuronen* entdeckt, deren Feuermuster mit bestimmten Positionen von Schallquellen korrespondierte. Im ebenfalls für die Lokalisation wichtigen Wo-Strom, der sich vom primären auditiven Kortex über den Gürtel bis zum erweiterten Gürtel erstreckt, sind die Neuronen entsprechend dieser Reihenfolge immer mehr auf die Positionen der Schallquellen, insbesondere auch von komplexen Klängen, abgestimmt [4, S. 295-301].

In der natürlichen Umwelt treten oft mehrere Schallereignisse an verschiedenen Orten gleichzeitig auf. Um die einzelnen Schallquellen voneinander zu trennen, muss diese als *auditive Szenerie* betrachtete Anordnung durch das auditive System mittels Analyse in separate Wahrnehmungen zerlegt werden. Für diese *auditive Gruppierung* nutzt das auditive System verschiedene Prinzipien, die auf Hinweisen bzw. Erfahrungen aus der Umwelt basieren. Da Schallereignisse von derselben räumlichen Position in der Regel von einer Schallquelle ausgehen, hilft die Trennung der *Herkunftsorte* mehrerer Schallquellen bei der getrennten Wahrnehmung der Schallereignisse. Ebenso stammen Schallereignisse mit *ähnlicher Klangfarbe* bzw. Spektrum oft von derselben Quelle, weshalb diese Stimuli vom auditiven System meist gruppiert werden. Dies wird insbesondere bei der *auditiven Sequenzgliederung* deutlich. Dabei werden die Schallereignisse entsprechend ihres Spektrums in unterschiedliche Wahrnehmungsströme separiert. So kommt es beispielsweise zu einer auditiven Sequenzgliederung, wenn sich tiefe und hohe Töne genügend schnell abwechseln. Statt einer alternierenden Melodie werden dann eine hohe und eine tiefe Melodie

wahrgenommen. Gleichzeitig tritt hier das Prinzip der *zeitlichen Nähe* auf, bei dem schnell aufeinander folgende Schallereignisse als in der Regel von derselben Schallquelle stammend, wahrgenommen werden. Auch konstante bzw. sich nur langsam in ihrer Frequenz ändernde Schallsignale mit *gutem Verlauf* werden -sogar bei Unterbrechungen- als kontinuierliche Stimuli derselben Schallquelle wahrgenommen. Letztlich fließt die Erfahrung jedes auditiven Systems in die Bewertung einer auditiven Szenerie mit ein. So werden Melodien in *Gedächtnisschemen* gespeichert. Diese stehen dann dem auditiven System zum Vergleich mit neu aufgenommenen Stimuli zur Verfügung. Die genannten Prinzipien funktionieren einzeln betrachtet jedoch nicht fehlerfrei, sodass es zu Täuschungen kommen kann. Nur unter Einbeziehung möglichst vieler dieser Prinzipien kann das auditive System Fehler in der Wahrnehmung der Umwelt vermeiden. Diese Strategie ähnelt der des visuellen Systems, das ebenfalls möglichst viele Organisationsprinzipien und insbesondere die Tiefenkriterien der Raumwahrnehmung nutzt (Vgl. Abschnitt 2.3., S. 11-12), um die Umwelt optimal wahrzunehmen und daraus die für das Überleben wichtigsten Entscheidungen bzw. Handlungen abzuleiten [4, S. 302-307].

3.3.5 Auditive Wahrnehmung in geschlossenen Räumen

Da im Rahmen des SEACEN-Teil-Projektes P9 der Einfluss akustischer und optischer Eigenschaften von Konzertsälen als Stimuli für die Raumwahrnehmung untersucht wird, soll im Folgenden die auditive Wahrnehmung in geschlossenen Räumen näher betrachtet werden.

Im Gegensatz zu dem im Freifeld¹ vorhandenen Direkt- bzw. *Primärschall* (Abbildung 67) treten in geschlossenen Räumen zahlreiche Reflexionen an Wänden, Decken, Boden etc. auf, die unsere Ohren nicht direkt sondern über den entsprechenden indirekten Weg erreichen (Abbildung 68). Das allseitige Auftreten der Reflexionen könnte für eine eindeutige Lokalisation der Schallquelle ein Problem darstellen. Jedoch wird auch in geschlossenen

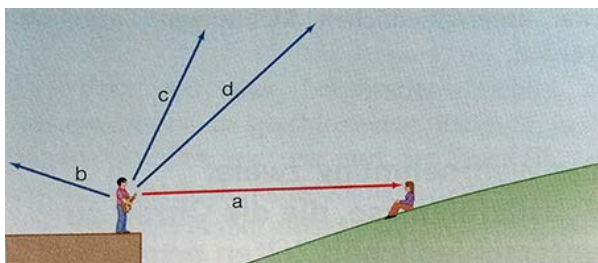


Abbildung 67: im Freifeld wird nur der Primärschall (a) wahrgenommen [4, Abb. S. 307]

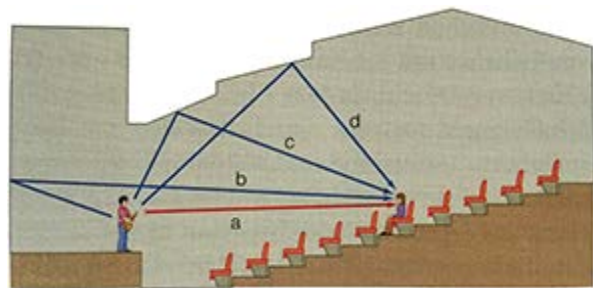


Abbildung 68: in geschlossenen Räumen werden sowohl der Primärschall (a) und der reflektierte Schall von (b) (c) und (d) wahrgenommen [4, Abb. S. 307]

¹ Umgebung ohne auftretende Reflexionen

Räumen in der Regel eine Schallquelle gut lokalisiert. Eine entscheidende Rolle spielt dabei die Verzögerungszeit zwischen dem Direktschall und den Reflexionen. Entsprechende Forschungen zeigen, dass es bei Verzögerungen unterhalb einer Millisekunde zur *Verschmelzung* zweier Schallsignale kommt, die dann als ein Schallsignal wahrgenommen werden [4, S. 307-308]. Liegt die Verzögerungszeit unterhalb der *Echoschwelle*, deren Wert zwischen einer und mehreren hundert Millisekunden betragen kann, wird das Gesamtsignal da geortet, wo sich die Schallquelle des zuerst eintreffenden Signals befindet (*Präzedenzeffekt*). Oberhalb dieses signal- und pegelabhängigen Wertes werden zwei separate Hörereignisse in Form eines Echos wahrgenommen (Abbildung 69). Wenn die Pegelminderung des Sekundärschalls beispielsweise 10dB beträgt, liegen typische Verzögerungszeiten für Sprache bei 50ms und für Musik bei 80ms [68, S. 103]. Die genauen Werte hängen vom Signal, vom Pegel und auch von den Dimensionen und Formen des Raumes ab.

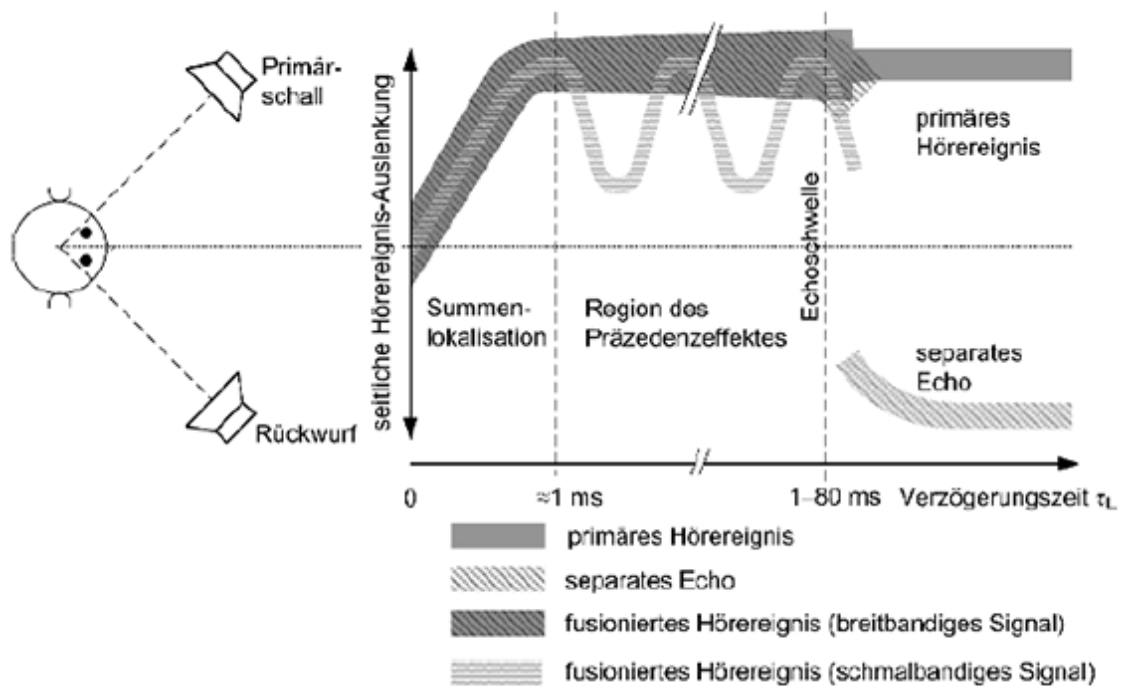


Abbildung 69: Lokalisation von Schallereignissen in Abhängigkeit von der Verzögerungszeit unter Einfluss des Präzedenzeffektes [68, Abb. S. 104].

Da im Rahmen des SEACEN-Projektes neben optischen Eigenschaften insbesondere der Einfluss akustischer Eigenschaften von Konzertsälen auf die auditive und visuelle Raumwahrnehmung untersucht werden soll, ist es für die Erzeugung der akustischen CAVE notwendig, eine möglichst genaue Akquise der akustischen Daten bzw. Parameter der Konzertsäle durchzuführen. Die grundlegende Technik für diese Datenakquise soll im nächsten Abschnitt erläutert werden.

3.4 Adaption der Technik an die auditive Wahrnehmung

Die im vorangegangenen Abschnitt dargestellten Eigenschaften des auditiven Systems bilden die Grundlage für das Verständnis der im Folgenden gezeigten Möglichkeiten der Erzeugung einer akustischen CAVE. Wie bei der Anpassung an die Eigenschaften der visuellen Wahrnehmung (Vgl. Abschnitt 3.2, S. 26) muss bei der Adaption an das auditive System ein akustisches Übertragungssystem geschaffen werden, das sowohl die Akquise geeigneter, dem Ohr äquivalent erfasster Audiodaten als auch die Wiedergabe dieser Daten in einer für die Ohren unverfälschten, plausiblen Form gewährleistet. Da die Nutzung der CAVE zeitunabhängig und wiederholbar erfolgen soll, ist die sichere (nach Möglichkeit redundante) und unkomprimierte Speicherung der Daten in einem geeigneten Datenformat erforderlich. Des Weiteren sollte eine schnelle, flexible Ausgabe der Daten möglich sein.

3.4.1 Akustisches Aufnahmesystem für eine CAVE

3.4.1.1 Mikrofonierung zur akustischen Akquise

Während beim visuellen System das Licht bzw. das von Objekten und der Umwelt reflektierte Lichtspektrum mittels Bildsensoren aufgezeichnet, gespeichert und weiterverarbeitet wird, erfassen Audioaufnahmesysteme den durch Druckunterschiede in der Materie entstehenden Schall.

Während Hydrofone in Flüssigkeiten und Körperschallmikrofone in Festkörpern Schall aufnehmen, erfolgt die Akquise von Audiodaten in der Luft mit Hilfe von *Mikrofonen* als Schallempfängern [69, S. 327]. Diese elektroakustischen Wandler erfassen den Luftschall und wandeln ihn in elektrische Signale um. Zunächst erfolgt dabei die Umwandlung der Schallwellen in eine mechanische Auslenkbewegung eines Festkörpers, meist einer Membran. Die dadurch entstehenden Membranschwingungen werden in elektrische Signale transformiert. Entsprechend ihren Eigenschaften werden Mikrofone nach bestimmten Kriterien klassifiziert (Abbildung 70). Da vom Mikrofon als erstem Übertragungselement in hohem Maße die Güte der gesamten Übertragungstrecke abhängt, muss bei der Auswahl der Mikrofone besonderen Wert auf höchste Qualität gelegt werden. Um im Rahmen des SEACEN-Projektes insbesondere die ästhetischen Eigenschaften des aufzunehmenden Streichquartetts zu erfassen, stehen die klangbildenden Eigenschaften der auszuwählenden Mikrofone im Vordergrund. Im Gegensatz dazu gelten für Messmikrofone im Wesentlichen technische Kriterien [70, S. 314]. In Abhängigkeit von der einwirkenden Kraft wird zwischen *Druckempfängern*, auf die der Schalldruck p (N/m^2) wirkt und *Druckgradientenempfängern*, auf die ein Schalldruckgefälle (Gradient) $\text{grad } p$ (N/m^3) einwirkt, unterschieden. Während bei letzteren Richtcharakteristiken von *Niere* über *Hyperniere* bis *Acht* realisiert werden können, tritt bei Druckempfängern ausschließlich eine Kugelcharakteristik auf [69, S. 328-330].

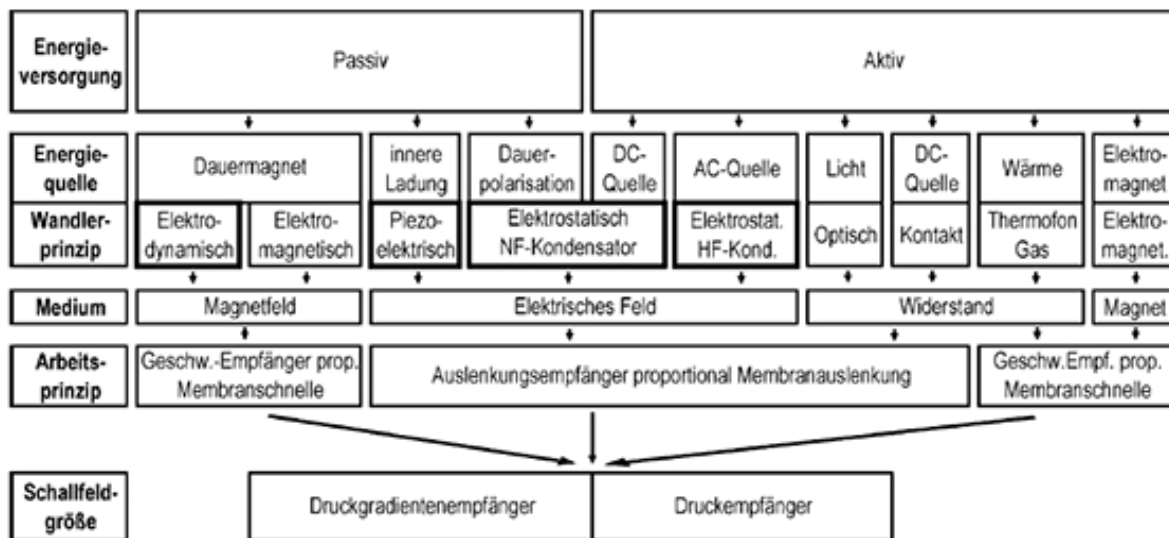


Abbildung 70: Klassifizierung von Luftschallmikrofonen nach verschiedenen Kriterien [70, Abb. S. 315]

Aufgrund der Vielzahl an Mikrofontypen werden nur die zwei wichtigsten für das SEACEN-Projekt in Frage kommenden Typen vorgestellt.

Elektrodynamische Mikrofone funktionieren nach dem Prinzip der elektromagnetischen Induktion. In einem senkrecht zum Leiter stehenden Magnetfeld bewegen die Schallwellen mechanisch einen Leiter, wodurch eine elektrische Spannung in diesem Leiter induziert wird (Abbildung 71). Diese Ausgangsspannung U_a ist zur magnetischen Flussdichte B , zur Länge des Leiters und zur Geschwindigkeit v der Membranbewegung proportional und spiegelt damit das akustische Signal der Schallwellen wider. Es gilt folgender Zusammenhang:

$U_a = B l v$	U_a = Ausgangsspannung B = magnetischen Flussdichte l = Länge des Leiters v = Geschwindigkeit der Membranbewegung
---------------	--

Die am weitesten verbreiteten elektrodynamischen Bauformen sind *Bändchenmikrofone* und *Tauchspulenmikrofone* [70, S. 320]. Die letzt genannte Gruppe soll wegen eines möglichen Einsatzes beim SEACEN-Projekt näher beschrieben werden.

Das Tauchspulenmikrofon besteht aus einem in Form einer Spule gewickelten, beweglichen Leiter, der mit einer Membran verbunden (verklebt) ist und sich im Magnetfeld eines topfförmigen Permanentmagneten befindet. Bei eintreffenden Schallwellen bewegt sich die Membran und damit der Leiter, wodurch eine Spannung im Leiter induziert wird, die zur Membrangeschwindigkeit und damit zu dem akustischen Signal proportional ist [70, S. 322].

Um einen möglichst linearen Frequenzgang zu erreichen, legt man bei Druckempfängern die stark zu bedämpfende Membranresonanz in die Mitte des gesamten Übertragungsbereiches. Durch zusätzliche, gedämpfte Resonatoren zu beiden Seiten der Systemresonanz kann der resultierende Frequenzgang begradigt und sogar spektral erweitert werden. Bei Druckgradientenempfängern erfolgt hingegen eine „Tiefenabstimmung“, d.h. die Systemresonanz

wird unterhalb des angestrebten spektralen Übertragungsbereiches verschoben. Dadurch kann ein ebener Frequenzgang erreicht werden [69, S. 331].

Elektrostatische Mikrofone nutzen die Spannungsänderung, die sich bei Abstands- und damit Kapazitätsänderung der Platten eines geladenen Kondensators ergibt (Abbildung 72). Der einfallende Schall bewegt dabei die elektrisch leitende Membran, die in geringem Abstand zu einer starren metallischen Gegenelektrode angeordnet ist. Es gilt folgender Zusammenhang:

$Q = C U = \epsilon \frac{S}{a} U$	Q = Kondensatorladung	S = Elektrodenfläche
	C = Kondensatorkapazität	d = Elektrodenabstand
	U = Kondensatorspannung	ϵ = Dielektrizitätskonstante

Die wichtigste Bauform sind *NF¹-Kondensatormikrofone*. Diese werden durch Anlegen einer Polarisationsspannung betrieben. Durch geringe Membranmasse und kurze Impulsantwort kann ein breiter Frequenzbereich und ein linearer Frequenzgang erzielt werden. Kondensatormikrofone besitzen von allen Wandlertypen den höchsten Dynamikumfang (ca. 130 dB) bei gleichzeitig geringster Geräuschspannung [70, S. 323-324].

Diese Eigenschaften bilden geeignete Voraussetzungen für den Einsatz von NF-Kondensatormikrofonen zur Audioaufnahme im SEACEN-Projekt.

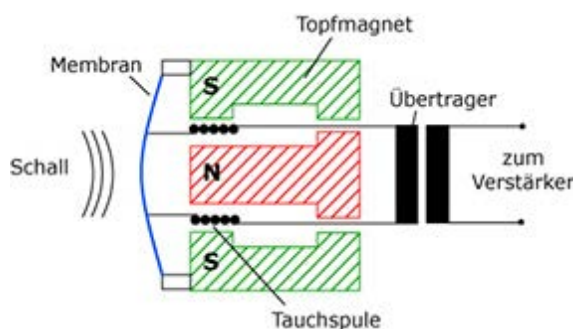


Abbildung 71: dynamisches Mikrofon, Prinzip [71]

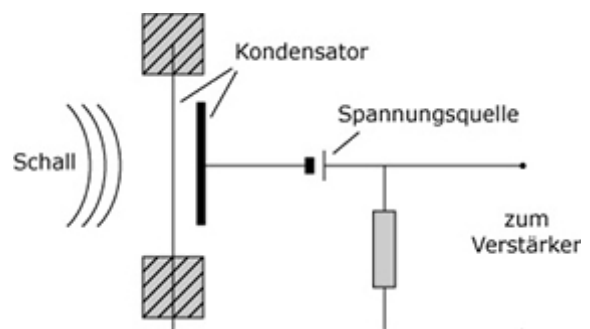


Abbildung 72: Kondensatormikrofon, Prinzip [72]

3.4.1.2 FABIAN – ein System zur automatisierten Akquise kopfbezogener BRIRs

Um eine Schallquelle, die sich in einem akustischen Raum mit spezifischen Halleigenschaften befindet, plausibel nachzubilden, ist es möglich, eine akustische Akquise der Schallquelle in einem nachhallfreien Aufnahmeraum mit der Impulsantwort des spezifischen akustischen Raumes zu verbinden. In Bezug auf das SEACEN-Projekt sollten Darbietungen eines Streichquartetts, die mittels geeigneter Mikrofone im hallfreien Raum aufgenommen wurden, später mit Hallräumen von sechs Konzertsälen in der CAVE wiedergegeben werden.

Während beim Ambisonics-Verfahren [Vgl. 73, S. 659] bzw. bei der Wellenfeldsynthese [Vgl. 73, S. 664] Schallfelder in dem betreffenden Raumvolumen erzeugt werden, kommt in der *Binauraltechnik* ein hörerzentriertes Verfahren zur Anwendung. Dieses Verfahren basiert auf

¹ NF – Niederfrequenz

der Nachbildung des Schalldruckverlaufes vor dem Trommelfell. Darin sind alle für die auditive Wahrnehmung wichtigen Informationen wie spektrale Zusammensetzung, Intensität und Lokalisation von Schallquellen enthalten und können bei der Wiedergabe die akustische Umgebung des Trommelfells rekonstruieren. Die Schallübertragung kann im Zeitbereich durch eine Impulsantwort beschrieben werden. Das resultierende binaurale Signal an beiden Trommelfellen ergibt sich somit aus der Faltung des nachhallfreien Signals vom Sendeort mit den beiden kopfbezogenen Raumimpulsantworten (Abbildung 73). Zur Aufnahme dieser BRIRs¹ werden hauptsächlich Kunstkopfsysteme eingesetzt, bei denen anstelle des Trommelfells Mikrofone zum Einsatz kommen. Die BRIRs hängen von den individuellen Kopf- und Außenohr-Dimensionen ab. Abweichungen können bei der akustischen Akquise durch geeignete Durchschnittsformen von Torso, Kopf und Ohr gering gehalten werden. Fehler können durch Einflüsse von Kopfhörer (bei der Wiedergabe) und Messmikrofonen entstehen, die durch entsprechende Filter beseitigt werden sollten. Bewegungen, insbesondere Drehungen des Kopfes der Versuchspersonen bei der Wiedergabe in der CAVE können nur durch abtastende Positionssensoren und die daraus resultierende korrigierte, verzögerungsfreie Impulsantwort ausgeglichen werden. Da die Wiedergabe kopfbezogener Signale über Lautsprecher problematisch ist, kommen beim SEACEN-Projekt Kopfhörer zum Einsatz. Der bei der Binauraltechnik möglicherweise verfälschende Einfluss insbesondere visueller Informationen auf die auditive Wahrnehmung ist u.a. Gegenstand der Untersuchungen im SEACEN-Projekt. [73, S. 671-672].

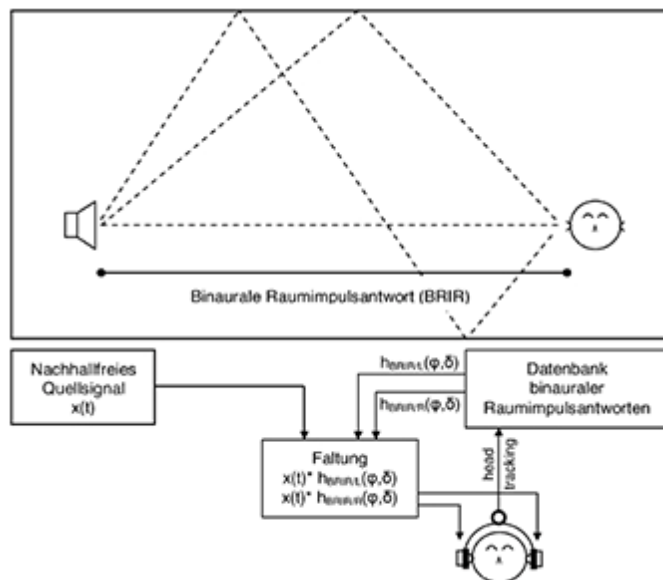


Abbildung 73: Binauraltechnik: Aufnahme der BRIRs mit Kunstkopf (oben) und Wiedergabe der Faltung aus nachhallfreiem Quellsignal und BRIRs über Kopfhörer (unten) [73, Abb. S. 672]



Abbildung 74: Messsystem FABIAN zur Akquise binauraler Raumimpulsantworten [74, Abb. S. 3]

¹ Binaural Room Impulse Response – binaurale Raumimpulsantwort

Das an der TU Berlin 2006 von Lindau entwickelte Messsystem FABIAN¹ ermöglicht die Akquise binauraler Impulsantworten mit Hilfe eines durch einen Schrittmotor betriebenen, hochpräzisen servoelektrischen Schwenk-Neige-Gelenks vom Typ Amtec Robotics PW-070. Dabei können Rotationsbewegungen von -180° bis $+180^\circ$, also ein vollständiger 360° -Schwenk mit einer minimalen Rotationsschrittweite von $0,1^\circ$ realisiert werden. [13, S. 35]. Die schnelle Positionierung des Kopfes erfolgt dabei mit einer Genauigkeit von $0,002^\circ$. Zur Schallaufnahme werden sehr rauscharme, in den Boden der Ohrmuschel eingelassene, körperschallentkoppelte Miniaturkondensatormikrofone vom Typ DPA 4060 (\varnothing 5.6 mm) eingesetzt. Die Steuerung erfolgt vollautomatisch mit Hilfe einer MatLab-basierten Software entsprechend eingegebener Parameter wie einem frei wählbaren Raster an Winkelschritten. Durch die verwendeten ASIO-Treiber² können Aufnahmen mit Sampelraten zwischen 44,1 und 96 kHz und einem 32 bit floating point Datenformat realisiert werden. Als Messsignal wird ein kontinuierlicher, frequenzgewichteter Sinussweep mit konstanter Amplitude verwendet [75, S. 1]. Dieser wird über Lautsprecher, die sich an der Position der Künstler als Ersatz der natürlichen Schallquellen befinden, wiedergegeben. Da Streichinstrumente nur eine geringe Richtwirkung besitzen und damit überwiegend allseitig abstrahlen, kommen Weitwinkel-Lautsprecher vom Typ QSC K8 zum Einsatz, die mit einem Öffnungswinkel von 105° konisch den natürlichen Schallquellen entsprechen. Diese Lautsprecher, deren Frequenzgang entzerrt wurde, erreichen einen Schalldruckpegel von 127 dB und besitzen einen breiten Frequenzbereich von 61 Hz - 20 kHz. Der lineare Sinussweep wird mit 44,1 kHz Abtastrate und 32 Bit Wortlänge abgespielt. Zur optimalen Messung erfolgt die Ausrichtung der Lautsprecher direkt in Richtung der im automatisierten, motorgesteuerten Kopf- und Rumpfsimulators (HATS³) FABIAN befindlichen In-Ear-Mikrofone. Um eine plausible Auralisation⁴ auch bei Kopfbewegungen zu ermöglichen, werden für jede Schallquelle verschiedene horizontale Kopforientierungen mit einer genauen Winkelauflösung von 1° -Schritten im Bereich von -80° bis $+80^\circ$ gemessen [76, S. 297]. Die Integration der relevanten Messdaten der BRIR-Aufnahme im Rahmen des SEACEN-Projektes und deren Wiedergabe in der TU-CAVE werden in Kapitel 4, Abschnitt 4.3, S. 71 bzw. in Kapitel 7, Abschnitt 7.1.1 beschrieben.

3.4.2 Akustisches Wiedergabesystem für eine CAVE

Im Folgenden wird ein akustisches Wiedergabesystem vorgestellt, das die besonderen Anforderungen in einer virtuellen Umgebung berücksichtigt. Um die mittels Mikrofonen und des Messsystems FABIAN akquirierten Impulsantworten und die im nachhallfreien Raum

¹ Fast and Automatic Binaural Impulse response Acquisition - automatisierte Akquise binauraler Impulsantworten

² ASIO - Audio Stream Input/Output - standardisiertes Audioübertragungsprotokoll

³ HATS - Head And Torso Simulator – Kopf- und Rumpfsimulator

⁴ Höreindruck

aufgenommenen Schallereignisse plausibel wiederzugeben, ist die Verwendung von speziellen Kopfhörern sinnvoll.

Bei der Wiedergabe von binauralen Signalen über Lautsprecher treten einige Probleme auf. Zunächst verändern weitere durch die Raumimpulsantwort des Abhörtraumes und durch den Einfallswinkel der Lautsprecher zum Kopf bedingte binaurale Informationen das ursprüngliche binaurale Signal. Des Weiteren kommt es bei dieser *transauralen* Lautsprecherübertragung zu einer Überlagerung der beiden ursprünglich eindeutig dem linken bzw. rechten Ohr zugeordneten Signale, was zum Übersprechen jedes Ohres durch den jeweils anderen Lautsprecher und damit zum Verlust der räumlichen Perspektive und zu spektralen Änderungen des Klangbildes führt [73, S. 677]. Aufgrund dieser unvermeidbaren Klangverfälschungen bei der Wiedergabe über Lautsprecher, rückt die Möglichkeit einer Kopfhörer-basierten Wiedergabe der binauralen Reize in den Mittelpunkt.

Kopfhörer und Lautsprecher funktionieren wie auch dynamische Mikrofone nach dem reversiblen Prinzip der elektromagnetischen Induktion (Vgl. Abschnitt 3.4.1.1, S. 58). Im Gegensatz zum Mikrophon erfolgt bei Kopfhörern und Lautsprechern jedoch die Umwandlung von elektrischer Energie in mechanische Energie. Fließt durch die im Magnetfeld befindliche, frei bewegliche, mit der Membran verbundene Spule eine Spannung, so bewegt sich die Spule und damit die Membran proportional zum Spannungsverlauf [77, S. 143-144]. Diese mechanischen Auslenkbewegungen der Membran werden dann als Schallwellen abgestrahlt. Ein geeignetes Kopfhörersystem muss für eine transparente binaurale Wiedergabe eine Anzahl von Kriterien erfüllen. Neben einer hohen spektralen Bandbreite mit einer möglichst linearen Übertragungsfunktion sowie einem hohen Signal-Rausch-Verhältnis (SNR¹) und ausreichender Übersprechdämpfung, sollte der Frequenzgang möglichst unabhängig vom Standort und von der Außenohr-Anatomie des Benutzers sein. Darüber hinaus sollte das System durch einen geringen Abstand zwischen Ohr und Kopfhörer die so genannte freifeldäquivalente Ohrkanalankopplung (FEC²) erfüllen. Die sich daraus ergebende Näherung zwischen Strahlungsimpedanz der Luft und akustischer Impedanz des Kopfhörers ermöglicht u.a. die Durchführung vergleichender und erweiterter Messungen bzw. Experimente zwischen Kopfhörersignal und einem realen Schallfeld, das außerhalb des Kopfhörers erzeugt wird. Für den Einsatz in einer virtuellen Umgebung sollte das Kopfhörersystem mit den dort verwendeten Komponenten wie 3D-Shutter-Brille, am Kopf montierten Displays oder Head-Tracking-Sensoren kombinierbar sein [78, S. 313].

Da derzeit die oben genannten Forderungen an binaurale Kopfhörer von kommerziellen Systemen nicht erfüllt werden, begann das Fachgebiet Audiokommunikation an der TU Berlin 2010 einen Prototypen eines binauralen Kopfhörers als wichtige Komponente für ein

¹ SNR - signal-to-noise ratio

² FEC - free air equivalent coupling

binaurales Referenzsystem zu entwickeln. Dazu wurde eine geschlossene Lautsprecherbox mit einem geeigneten und leicht verfügbaren elektrodynamischen Miniatur Lautsprecher entworfen, der eine optimale Kombination aus Frequenzbandbreite, akustischer Leistung, Richtcharakteristik, Größe, Masse und Volumen darstellt. Die Zielvorgaben für das maximale Volumen von 200 ml bei einem Frequenzgang von 200 Hz - 15 kHz wurden vom Lautsprecher-Modell Tymphany Peerless 30970 bzw. 830983 erfüllt, der aufgrund dessen für den binauralen Kopfhörer zum Einsatz kam. Der mittels 3D-CAD Software modellierte Prototyp BK109¹ wiegt einschließlich Verdrahtung und Lautsprecher ca. 600 g (Abbildung 75). Der BK109 verfügt über einen linearen Vollbereichs-Nahfeld-Frequenzgang mit tolerierbaren nichtlinearen Verzerrungen bei mittlerer Lautstärke (Abbildung 76). Bei niedrigen Frequenzen mit hohen Schalldruckpegeln kann der Klirrfaktor² unter Umständen ein kritisches Maß erreichen. Das Übersprechen zwischen linkem und rechtem Kanal ist bei hohen Frequenzen vernachlässigbar und bei niedrigen Frequenzen ausreichend klein. Die freifeldäquivalente Ohrkanalankopplung FEC (Vgl. S 62) wird vom BK 109 erfüllt. Durch genaue Positionierung der Kopfhörer können individuelle anatomische Unterschiede der Benutzer minimiert werden. Optimiert werden müssen jedoch noch die Toleranz gegenüber mechanischen Schwingungen des Kopfhörer-Gehäuses und die Gestaltung einer zweckmäßigen anwenderfreundlichen mechanischen Befestigung. [79, S. 704-713].



Abbildung 75: durch einen Gurt an FABIAN befestigter binauraler Kopfhörer Prototyp BK109 [79, Abb. S. 704]

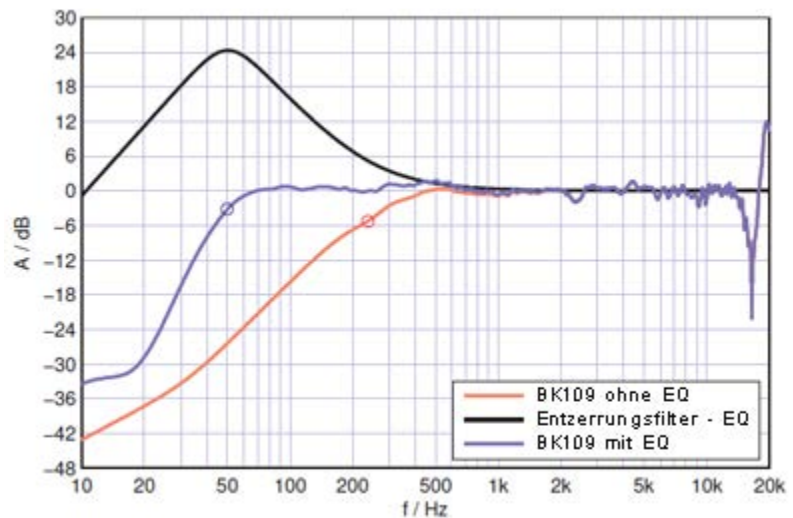


Abbildung 76: Nahfeld - Frequenzgang des binauralen Kopfhörers BK109 mit und ohne Entzerrungsfilter [79, Vgl. Abb. S. 706]

¹ BK – binauraler Kopfhörer

² Maß für nichtlineare Verzerrungen

4 Nachhallfreie Audioproduktion eines Streichquartetts

Nachdem die grundlegende technische Realisation einer virtuellen Umgebung eingehend erläutert wurde, werden in den nächsten Kapiteln die nachhallfreie Audioproduktion und die stereoskopische Videoproduktion/-postproduktion eines Streichquartetts dokumentiert. Die hier im Fokus stehenden akustischen wie auch optischen künstlerischen Inhalte sollen im Rahmen des SEACEN-Projektes P9 in eine virtuelle Umgebung integriert werden. Für eine plausible Raumwahrnehmung in der CAVE ist dabei die Kombination der nachhallfreien Audioinhalte mit der mittels FABIAN akquirierten Raumakustik (BRIRs) von sechs verschiedenen Konzertsälen und die Verbindung des im Vordergrund befindlichen stereoskopischen Streichquartetts mittels Greenscreentechnik mit sechs verschiedenen stereoskopischen Konzertsaalpanoramen erforderlich (Abbildung 77). In der Weiterführung des SEACEN-Projektes soll anschließend das Zusammenwirken unterschiedlicher, variierbarer optischer und akustischer Stimuli und deren Auswirkungen auf die visuelle und auditive Raumwahrnehmung untersucht werden.

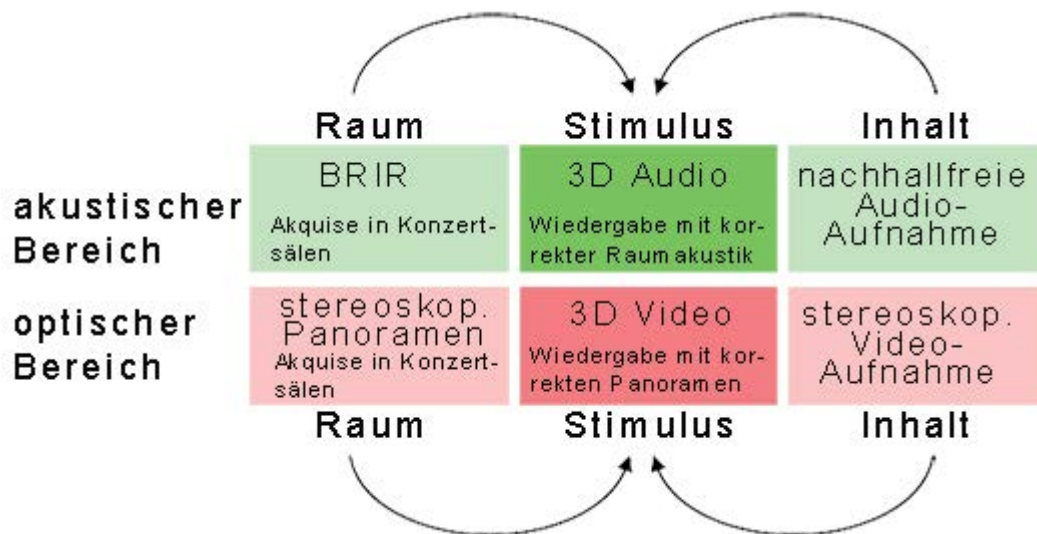


Abbildung 77: Konzept der Reizerzeugung für die SEACEN-CAVE; durch unabhängige Variationen von optischem und akustischem Bereich sowie von Raum und Inhalt können Zusammenhänge zwischen visueller und auditiver Raumwahrnehmung untersucht werden [76, Vgl. Abb. S. 294]

4.1 Vorbereitung

Um die spezifischen Eigenschaften der Konzerträume in der CAVE erfahrbar zu machen, wurden die Inhalte nach bestimmten Kriterien ausgewählt. Im akustischen Bereich zählte insbesondere die Impulshaftigkeit und die spektrale Breite des Anregungssignals dazu sowie ein hoher Dynamikumfang und Pausen, die den Nachhall eines Raumes leichter erfassen lassen. Um Versuchspersonen auch ohne besondere Kenntnisse in Epoche und Gattung des dargebotenen Werkes einen opportunen, emotionalen Zugang zu ermöglichen, wurde die

Werkwahl zeitlich auf das 18. Jh. bis zur ersten Hälfte des 20. Jh. und lokal auf Europa eingegrenzt. Damit die Anforderungen einschließlich einer hohen emotionalen Bandbreite erfüllt werden konnten, wurde ein eingespieltes Streichquartett ausgebildeter und berufserfahrener Künstler engagiert – das Berlin-Budapest Quartett bestehend aus: Dea Szücs (1. Violine – VI1), Éva Csermák (2. Violine -VI2), Itamar Ringel (Bratsche -VIa) und Ditta Rohmann (Cello -Vc). Als Musik wurde das vom Streichquartett bereits zum festen Repertoire gehörende Werk, 2. Satz für Streichquartett g-moll, op. 10, 1893, von Claude Debussy, ausgewählt.

Nach der Wahl der Musik und der Interpreten, musste ein Ort mit der Möglichkeit einer nachhallfreien Audioaufnahme gefunden werden. Da das SEACEN-Projekt P9 an der TU Berlin durchgeführt wird, war es naheliegend für diese Aufnahmen den reflexionsarmen Raum des Instituts für Technische Akustik an der TU Berlin zu nutzen, da es sonst kaum Orte mit Freifeldbedingungen, also Orte ohne Reflexionen gibt, an denen eine komplexe Audioaufnahme durchgeführt werden kann. Der Begriff „reflexionsarm“ beschreibt im Gegensatz zu „nachhallfrei“, dass selbst in solchen stark absorbierenden speziellen Aufnahme- bzw. Messräumen -wenn auch nur sehr geringe- Reflexionen auftreten. Da diese Reflexionen für die vorliegende Arbeit jedoch zu vernachlässigen sind, wird im weiteren Verlauf von „nachhallfreier“ Audioproduktion bzw. -aufnahme gesprochen.

Der reflexionsarme Raum der TU Berlin hat ein Volumen von 1070 m^3 und einen begehbaren Teil von 126 m^3 . Die untere Grenzfrequenz beträgt 63 Hz. Der mit einem begehbaren Netz ausgerüstete Raum absorbiert den Schall in allen Richtungen (Abbildung 78). Er ist akustisch vollständig vom Kontrollraum entkoppelt [Vgl. 80].

Um die vier Musiker des Streichquartetts akustisch zu separieren, wurden sie möglichst weit auseinander in vier Ecken gesetzt, und es wurden Stellwände zwischen ihnen platziert, die den direkten Schallweg zwischen den Instrumenten unterbrechen sollten. Andernfalls würde es zu Übersprechungen kommen, sodass die Audiosignale des einen Musikers zusammen in einem Kanal mit den Signalen eines anderen Musikers mit aufgenommen würden. Zur Vermeidung von Reflexionen waren die Stellwände mit Decken behängt (Abbildung 79).



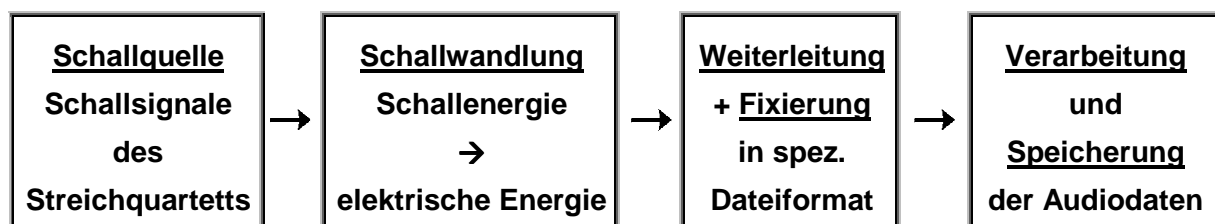
Abbildung 78: reflexionsarmer Raum des Instituts für Technische Akustik an der TU Berlin [80, Abb.].



Abbildung 79: Separierung der Musiker durch decken-behängte Stellwände im reflexionsarmen Raum

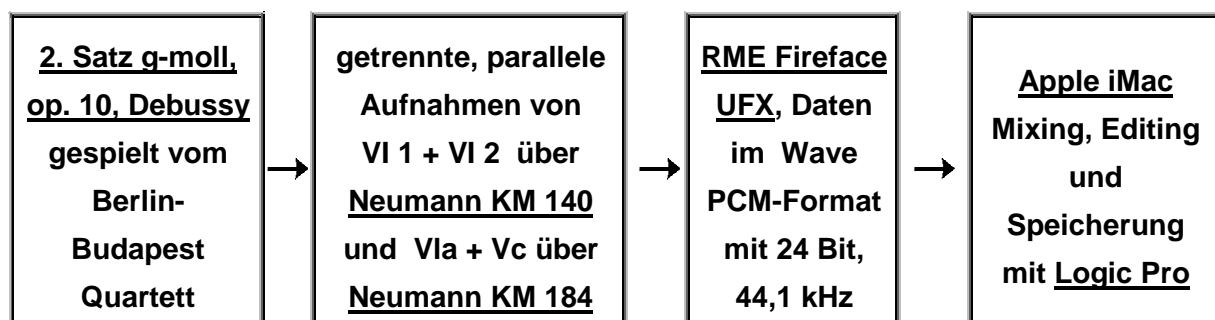
Für eine hochwertige, in der CAVE plausibel wirkende Audiowiedergabe ist der Einsatz entsprechender Technik bei der Audioaufnahme unabdingbar. Deshalb musste im Vorfeld der Produktion das passende Equipment akquiriert werden. Nach Aufstellung einer Equipmentliste konnten die meisten Komponenten vom Fachgebiet Audiokommunikation am Institut für Sprache und Kommunikation der TU Berlin geliehen werden. Darüber hinaus wurden acht Stellwände und ein Kopfhörerverstärker von der Universität der Künste (UdK) entliehen (Anhang A1).

Für die optimale Aufnahme des Streichquartetts im reflexionsarmen Raum und die Weiterleitung, Verarbeitung und Speicherung der entstandenen digitalen Audiodaten wurde eine Verarbeitungskette entworfen, die in folgendem Flussdiagramm dargestellt ist:



4.2 Durchführung

Die dreistündigen Audioaufnahmen im Rahmen des SEACEN-Projektes P9 wurden am 06.11.2011 im reflexionsarmen Raum des Instituts für Technische Akustik an der TU Berlin unter der Leitung von Dr. Hans-Joachim Maempel durchgeführt. Nach Auswahl aller Komponenten für die Audioaufnahmen entstand folgendes Flussdiagramm:



Um einen reibungslosen Ablauf der Audioaufnahmen zu gewährleisten, mussten neben dem Funktionieren der Technik insbesondere die Bedürfnisse der Musiker berücksichtigt werden. Dazu gehörte auch das Hinzufügen eines Halleffektes über Kopfhörer als Ersatz für das natürliche gegenseitige Hören der Musiker untereinander. Wichtig für eine authentische Aufnahme ist insbesondere die Wahl der passenden Mikrofone. Als Schallwandler kamen die verfügbaren Kleinmembran-Kondensatormikrofone Neumann KM 140 bzw. Neumann KM 184 mit Nierencharakteristik zum Einsatz. Die beiden Mikrofon-Typen unterscheiden sich nur geringfügig in ihren akustischen Eigenschaften und technischen Daten (Anhang A2). Sie besitzen einen weitgehend linearen Frequenzgang mit einer Erhöhung um circa 3 dB bei 8

kHz. Das Polardiagramm für die Abbildung der Richtcharakteristik des KM 140 weist bei den tiefsten und höchsten Frequenzen einen etwas unruhigen Verlauf auf [Vgl. 81, S. 68]. Insgesamt bieten beide Mikrofontypen eine sehr gute Nierencharakteristik. Die ideal für höherfrequente Instrumente geeigneten Mikrofone vom Typ KM 140 [Vgl. 81, S. 73] wurden für die Violinen verwendet. Für die Aufnahmen von Bratsche und Cello kam jeweils ein KM 184 zur Anwendung. Damit bei den Mikrofonaufnahmen ein ausgewogener Frequenzgang erreicht und gleichzeitig der Nahbesprechungseffekt¹ verhindert wird, musste ein ausreichender Abstand von dem jeweiligen Instrument eingehalten werden. Zu groß durfte der Abstand jedoch nicht gewählt werden, damit der durch die Stellwände nicht vollständig absorbierte Klang der anderen Instrumente hinreichend klein gegenüber dem aufzunehmenden Instrument blieb (Abbildung 81). Bei der Mikrofonausrichtung mussten sowohl der Schalleinfallswinkel bezogen auf das Mikrofon als auch die Abstrahlcharakteristik der Instrumente berücksichtigt werden.

Da alle Mikrofone eine Richtwirkung zu hohen Frequenzen hin aufweisen, tritt bei Abweichungen von der frontalen Schalleinfallrichtung (Einfallswinkel von 0° zur Mikrofonhauptachse) ein Höhenabfall auf. Je größer der Schalleinfallswinkel ist, desto mehr hohe Frequenzanteile gehen verloren (Abbildung 80). Um diesen Höhenabfall zu vermeiden und einen linearen Frequenzgang zu gewährleisten, sollte die Mikrofonachse senkrecht auf die Abstrahl-Ebene des Instrumenten-Resonanzraumes treffen, sodass der Schalleinfallswinkel 0° beträgt. Einen weiteren Einfluss auf den Frequenzgang hat die Größe der Mikrofonmembran. Bei Kleinmembran-Mikrofonen ist der Höhenabfall infolge der Abweichungen von der frontalen Schalleinfallrichtung nicht so stark wie bei Großmembran-Mikrofonen, was auch für den Einsatz von Kleinmembran-Mikrofonen spricht [82, S. 570-571].

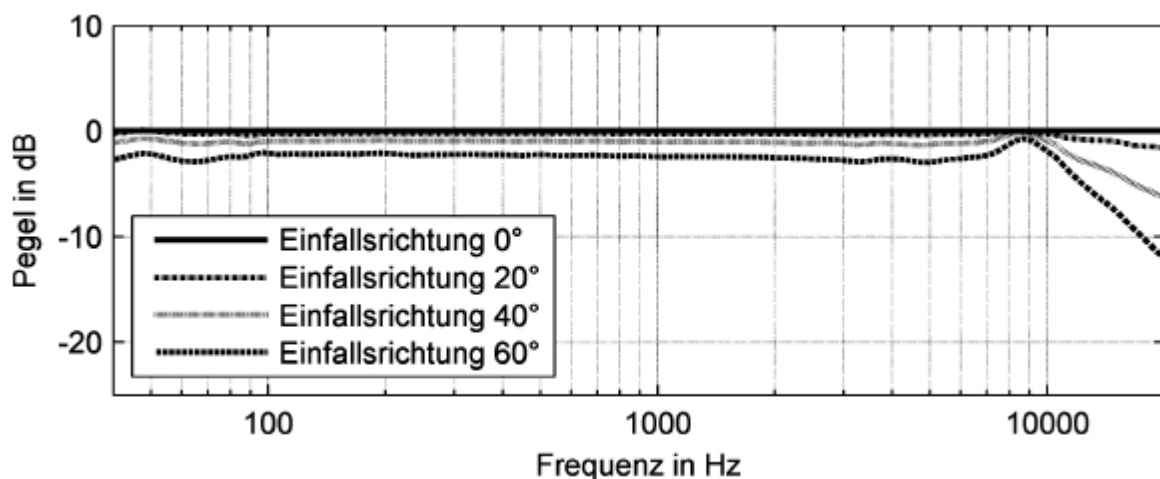


Abbildung 80: Frequenzgänge des für Bratsche und Cello im reflexionsarmen Raum (Freifeldbedingungen) eingesetzten Aufnahmемikrofons KM 184 in Abhängigkeit von der Schalleinfallrichtung [82, Abb. S.571]

¹ überproportionale Betonung der tiefen Frequenzen

Die Abstrahlcharakteristik von Streichinstrumenten ist u.a. durch ihren inneren asymmetrischen Aufbau von einer unregelmäßigen, unsymmetrischen Schallabstrahlung gekennzeichnet. Bei Violinen, die einen Frequenzumfang von 200 Hz bis 10 kHz besitzen, werden tiefe Frequenzen bis 500 Hz allseitig abgestrahlt. Höhere Frequenzen fokussieren sich dagegen in einer Hauptabstrahlrichtung senkrecht zum Korpus¹ (Abbildung 82, 83 rot) in Höhe der f-Löcher² innerhalb eines nur sehr kleinen Abstrahlwinkels von ca. 15° [83, S. 117]. Die Justierung des Mikrofon-schalleinfallswinkels, der zwischen der Mikrofonhauptachse und der Verbindung vom Mikrofonkopf zum jeweiligen Instrumenten-Hauptabstrahlpunkt lag, erfolgte entsprechend den Kriterien einer optimalen Klangwiedergabe des jeweiligen Instrumentes (Abbildung 82, 83, Anhang A3). Die Instrumentenabstrahlwinkel, die zwischen der Hauptabstrahlrichtung der Instrumente und der Verbindung vom Mikrofonkopf zum jeweiligen Instrumenten-Hauptabstrahlpunkt lagen (Abbildung 82, 83 grün) und die Mikrofon-schalleinfallswinkel (Abbildung 82, 83 gelb) wurden nach den Audioaufnahmen im reflexionsarmen Raum grafisch durch das Einzeichnen in entsprechende Fotos von vorn und von der Seite jedes Musikers mit Instrument ermittelt. So betragen die Instrumenten-abstrahlwinkel für alle Instrumente zwischen 2° und 13° und lagen damit innerhalb des 15°-Grenzwertes, der auch für die Bratsche und das Violoncello anwendbar ist [Vgl. 83, S. 119-120]. Das bedeutet, dass das gesamte Frequenzspektrum der Instrumente bei der Aufnahme erfasst wurde. Die Mikrofon-schalleinfallswinkel lagen für alle Instrumente zwischen 3° und 17° und wiesen damit nach Abbildung 80 nur einen geringen Höhenabfall entsprechend der Einfallrichtung 20° auf, und trugen so zur Erzeugung eines ausgewogenen Klangbildes bei.

Instrument	Mikrofon-Instrumenten Abstand in cm	Mikrofon-schalleinfallswinkel (gelb)		Instrumentenabstrahlwinkel (grün)	
		Winkel in ° vorn	Winkel in ° Seite	Winkel in ° vorn	Winkel in ° Seite
Violine 1 VI 1	66	11	17	11	5
Violine 2 VI 2	87	17	6	13	3
Bratsche Vla	88	5	3	2	4
Cello Vc	62	4 (hinten)	12	13 (hinten)	12

Abbildung 81: Mikrofon-Instrumenten-Abstände, Mikrofon-schalleinfallswinkel und Instrumentenabstrahlwinkel

Damit sich die Musiker auf ein virtuoseres, emotionales Spielen ihrer Instrumente konzentrieren konnten, musste ihnen ein angenehmes Klangbild über Kopfhörer dargeboten werden. Zu diesem Zweck wurden ihnen Kopfhörer vom Typ Audio-Technica ATH M40fs zur Verfügung gestellt. Mit Hilfe des Kopfhörerverstärkers Behringer Powerplay Pro HA-4600 konnten sie neben der Gesamtlautstärke die individuelle Lautstärke ihres Instrumentes in einem Gesamtmix aller von den Musikern gespielten Instrumente regeln (Abbildung 84).

¹ Klangkörper der Violine

² f-förmige Resonanzöffnungen im Korpus der Violine

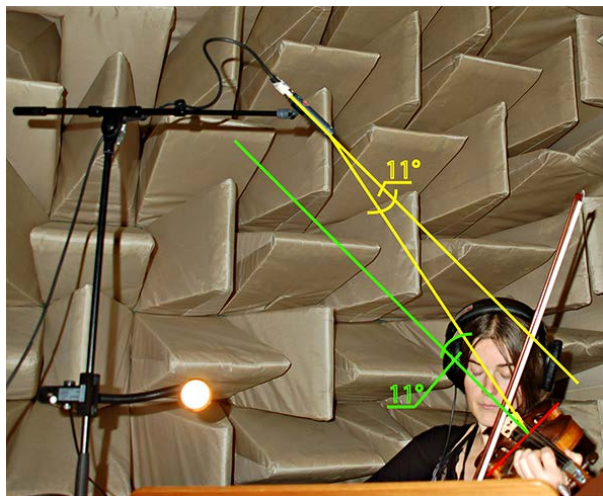


Abbildung 82: Mikrofonierung von Violine 1 (VI 1) mit passenden Abstands- und Winkeleinstellungen; vorn



Abbildung 83: Mikrofonierung von Violine 1 (VI 1) mit passenden Abstands- und Winkeleinstellungen; Seite

Neben der Verkabelung und Justierung der Mikrofone und Kopfhörer im reflexionsarmen Raum musste die Aufnahmebereitschaft des akustisch entkoppelten Kontrollraumes gewährleistet werden. Voraussetzung für die Weiterleitung der Audiosignale in ein Aufnahme- und Speichersystem war der Einsatz eines Multicore-Kabels, das durch einen Kanal vom reflexionsarmen Raum in den Kontrollraum geleitet wurde. Dort erfolgte die Audioaufnahme, die Abhörkontrolle der Audioaufnahmen, die Kommunikation mit den Musikern und deren Anleitung durch den Projektleiter Dr. Maempel. Die Assistenz wurde von Frau Mina Fallahi und dem Autor übernommen.

Das Abhören im Kontrollraum konnte durch den Einsatz der Studiomonitore Genelec 1031A realisiert werden. Dieses Zwei-Wege Aktiv Monitor System bestehend aus Bassreflexgehäuse und Verstärkereinheit mit aktiver Frequenzweiche bietet eine 120 W hohe Leistung und eine große Bandbreite von 47 Hz bis 22 kHz. Mittels der Directivity Control Waveguide¹ (DCW) Technologie kann ein exzellenter Stereosound und eine Anpassung des Lautsprechersystems an schwierige akustische Umgebungen durch vielseitige Kontrollmöglichkeiten erreicht werden [84]. Durch das ebenfalls am Regieplatz befindliche Talkback-Mikrofon Sennheiser Evolution E840 konnte mit den Musikern kommuniziert werden (Abbildung 85).

Da es für die Postproduktion und die damit verbundene Faltung der vier Streichinstrumente mit den BRIRs erforderlich war, jedes Instrument separat auf einer monofonen² Spur

¹ Steuerung der Richtcharakteristik

² einkanaliges Auftreten von Schallsignalen

aufzunehmen (Vgl. Abschnitt 4.3, S. 71), kam für die Aufnahme der Signale die Audioschnittstelle RME Fireface UFX zum Einsatz.

Dieses professionelle Audiointerface¹ bietet neben 60 Audiokanälen (davon zwölf analoge Ein- und Ausgänge) vier Mikrofonvorverstärker für die parallele Aufnahme der vier Streichinstrumente, Sampleraten bis zu 192 kHz, eine schnelle Datenübertragung über die FireWire 400-Schnittstelle sowie eine sehr geringe Audio-Latenz für verzögerungsfreies Aufnehmen und Monitoring. Der Hersteller spricht dabei von einer unverfälschten, originalgetreuen Klangübertragung mit außergewöhnlich niedrigen Rausch- und Verzerrungswerten [85].



Abbildung 84: Kopfhörerverstärker Behringer Powerplay Pro HA-4600 zum individuellen Kopfhörer-Mix für jeden Künstler



Abbildung 85: Regieplatz im Kontrollraum zum Abhören der Aufnahmen und Kommunizieren mit den Künstlern im reflexionsarmen Raum über Talkback

Entsprechend den Regieanweisungen durch Dr. Maempel wurden nach einer kurzen Einspielphase der Musiker so viele Takes² hintereinander aufgenommen, bis alle Musiker gleichzeitig ein fehlerfreies und emotional überzeugendes Spiel darboten.

Der im Flussdiagramm (Vgl. S. 66) dargestellte Signalfluss zeigt die Umwandlung des Instrumentenschalls in Audiodaten, deren Weiterleitung und Speicherung auf einer Computerfestplatte. Die durch die Mikrofone in elektrische Spannungen umgewandelten Instrumentenaufnahmen gelangten über das Multicore-Kabel in das RME Fireface UFX, das sich im vom reflexionsarmen Raum akustisch getrennten Kontrollraum befand. Nach Verstärkung im Fireface wurden die Audiosignale über die FireWire 400-Schnittstelle zu einem iMac Computer weitergeleitet. Dort erfolgt die softwaremäßige Aufnahme, Speicherung und Weiterverarbeitung der nun im Wave PCM-Format mit 24 Bit und 44,1 kHz vorliegenden Audiodaten mittels der Digital-Audio-Workstation-Software Logic Pro 9. Die Kontrolle und Steuerung der Aufnahmesession erfolgte über den Mixer Phonic AM8440.

¹ Audioschnittstelle

² Aufnahme eines vorläufigen Musik-Teilstückes oder eines gesamten Stückes

Das komplexe Routing aller Komponenten ermöglichte neben der parallelen Audioaufnahme aller vier Streichinstrumente auch die Anleitung der Künstler durch den Aufnahmeleiter mit Hilfe einer Verbindung zwischen Talkback-Mikrofon und Kopfhörer der Musiker. Die Musiker konnten wiederum über ihre Mikrofone vom Team im Kontrollraum gehört werden. Gleichzeitig wurden die Aufnahmen über die Boxen vom Typ Genelec 1031A abgehört, beurteilt und im Zweifelsfall wiederholt aufgenommen (Abbildung 85). In Abbildung 86 ist der technische Aufbau der nachhallfreien Audioproduktion dargestellt:

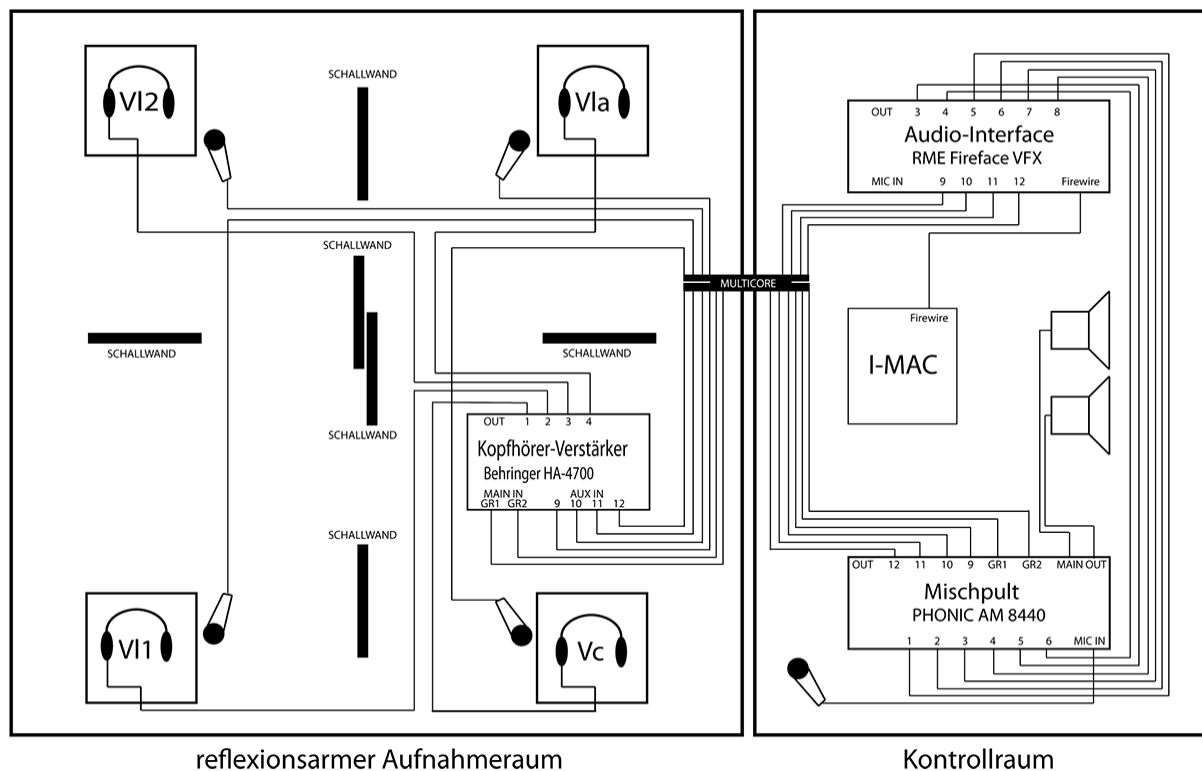


Abbildung 86: Technischer Aufbau der nachhallfreien Audioproduktion (ganzseitig siehe Anlage A4)

4.3 Postproduktion

Um die Authentizität des Live-Ereignisses zu erhalten, wurden die Aufnahmen nicht übermäßig bearbeitet. Einige wenige Spielfehler und grobe Unsauberkeiten wurden in der Postproduktion schnitttechnisch mit Hilfe des Editing-Programms Samplitude Music Studio 17 eliminiert [Vgl. 76, S. 300].

Aus den Audioaufnahmen im reflexionsarmen Raum resultierte eine Vier-Spur Aufnahme die eine monofone Spur jedes der vier Streichinstrumente beinhaltet. Diese enthielt jedoch noch keine Raum- bzw. Hallinformationen und keine Informationen über die Position jedes Instrumentes im Raum. Erst durch die Rechenoperation der Faltung der vier Spuren mit den BRIRs der sechs Konzertsäle wurde die Rauminformation jedes Konzertsalles hinzugefügt. Um die Faltung mittels moderner Digitaltechnik durchführen zu können, wurde das zeitkontinuierliche Faltungsintegral in eine zeitdiskrete Faltungssumme umgeformt. Durch eine hohe Rechenleistung lässt sich diese Operation mit dem Resultat einer genügend

großen Anzahl von Werten zu diskreten Zeitpunkten berechnen [Vgl. 86, S. 846]:

Faltungsintegral (zeitkontinuierlich)

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau) \cdot h(t-\tau) \quad \Rightarrow$$

* - Symbol für Faltungsoperation

y(t) - kontinuierliches Gesamtsignal mit Nachhall

x(t) – kontinuierliches nachhallfreies Signal

h(t) – kontinuierliche BRIR

zeitdiskrete Faltungssumme

$$y(n) = x(n) * h(n) = \sum_{k=-\infty}^{\infty} x(k) \cdot h(n-k)$$

* - Symbol für Faltungsoperation

y(n) – diskretes Gesamtsignal mit Nachhall

x(n) – diskretes nachhallfreies Signal

h(n) – diskrete BRIR

[Vgl. 87, S. 8-9, 18-19]

Während der BRIR-Akquise gelangten die von den Lautsprechern wiedergegebenen Messimpulse zum Kunstkopf des FABIAN-Messsystems (Abbildung 74). Jeder Lautsprecher repräsentierte dabei jeweils ein Instrument und war an dessen entsprechende Stelle platziert worden (Abbildung 87). Aus der computergestützten Berechnung der Faltung von BRIRs mit den nachhallfreien Aufnahmen resultierte eine Datei, die bei einer Wiedergabe über binaurale Kopfhörer die aurale Szenerie des jeweiligen Konzertsaaes mit dem spezifischen Nachhall bzw. den Rauminformationen und den exakten Positionen aller Instrumente plausibel widerpiegelte. Darüber hinaus ermöglichte der drehbar gelagerte Kopf des FABIAN-Messsystems die BRIR-Akquise bei verschiedenen Kopforientierungen. Dadurch ist bei hinreichend schneller Datenübertragung die Anwendung von Echtzeit-Head-Tracking¹ möglich. Ein Proband, der sich in der Cave befindet, erhält über Kopfhörer auch bei Kopfbewegungen innerhalb eines Grenzwinkels sich permanent anpassende binaurale Signale und damit einen plausiblen Raumeindruck (Vgl. Abschnitt 3.4.1.2, S. 59-61, Abbildung 73, 74).

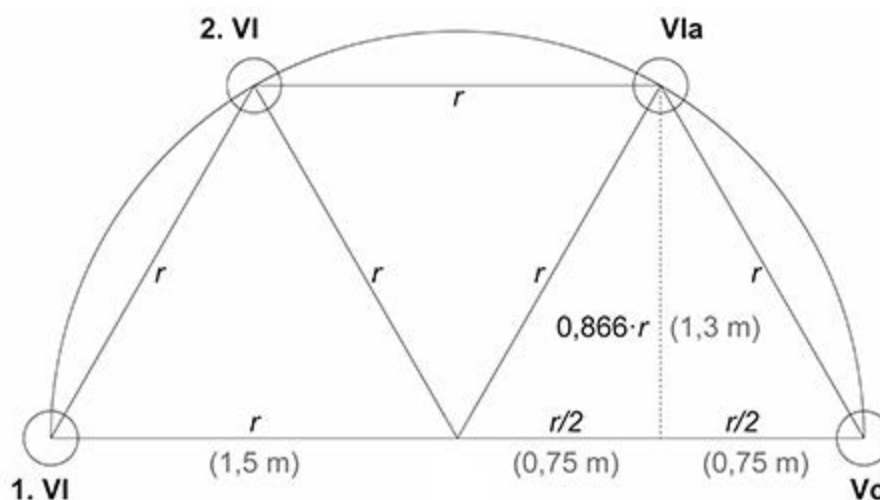


Abbildung 87: Positionen der einzelnen Streichinstrumente bei der Akquise der BRIRs [Vgl. 76, Abb. S. 296]

¹ Erfassung der Kopfbewegungen und daraus resultierende Anpassung eines Bewegungs-Folge-Systems

4.4 Fehler und Hemmnisse bei nachhallfreien Audioproduktionen

Das Ziel der Audioproduktion im reflexionsarmen Raum war die Erzeugung einer möglichst nachhallfreien Aufnahme des Streichquartetts. Aufgrund der großen Absorptionsfähigkeit des durch pyramidenförmige Schaumstoffe ausgekleideten reflexionsarmen Raumes und der Abschirmung der Instrumente gegeneinander durch Stellwände klingt die Aufnahme akustisch trocken, also nahezu nachhallfrei. Eine vollständige Eliminierung von Reflexionen bzw. des Übersprechens war jedoch nicht möglich, da die in Abbildung 86 maßstabsgerecht eingezeichneten Schallwände zu klein waren und auch keine idealen Absorber darstellten. Dennoch waren die Aufnahmen hinreichend nachhallfrei um Verfälschungen bei der Faltung mit den BRIRs zu vermeiden.

Ausführlich wurde in Abschnitt 4.2, S. 68 die Justierung des Mikrofonschalleinfallswinkels diskutiert. Abweichungen von der eingestellten Mikrofonausrichtung konnten durch die Veränderung der Position der Musiker während der Audioaufnahmen entstehen. Des Weiteren verändern sich durch emotionale und spieltechnische Bewegungen beim Spiel der Musiker permanent der Abstand und der Mikrofonschalleinfallswinkel wodurch das Klangspektrum wenn auch geringfügig beeinflusst wurde.

Ein technisch vermeidbares Hemmnis war der aus Kostengründen nicht beheizte reflexionsarme Raum. Die daraus resultierende Kälte war kontraproduktiv und behinderte z.T. ein sauberes Spiel der Musiker.

Ungeachtet der genannten Probleme und Hemmnisse waren die Resultate der Audioproduktion und der anschließenden Faltung mit den BRIRs geeignet, um in der TU-CAVE als akustische Stimuli eingesetzt zu werden.

5 Stereoskopische Videoproduktion und Postproduktion eines Streichquartetts

Nachdem im letzten Kapitel die Akquise akustischer Inhalte mit Hilfe der nachhallfreien Audioproduktion dargestellt wurde, wird im Folgenden die stereoskopische Videoproduktion, also die Akquise optischer Inhalte für eine virtuelle Umgebung erläutert.

Während für den Audibereich die BRIRs die akustischen Raumeigenschaften der sechs Konzertsäle darstellten, wurden die Raumeigenschaften für den optischen Bereich durch die stereoskopischen Hintergrund-Panoramen der Konzertsäle repräsentiert. Das Äquivalent zur Aufnahme im reflexionsarmen Raum im akustischen Bereich stellte im optischen Bereich die stereoskopische Videoaufnahme des Streichquartetts dar. Diese Streichquartett-Akquise ohne Hintergrund, also raumunabhängig, sollte durch eine in diesem Kapitel dokumentierte stereoskopische Greenscreen-Aufnahme des Streichquartetts erfolgen (Abbildung 88).

Um diese Aufnahmen durchführen zu können war neben einer genauen Planung und der Koordination aller Beteiligten, einer Materiallistenstellung und der Aufarbeitung der schon in Abschnitt 3.2.3 S. 40 ff dargelegten Grundlagen über stereoskopische Filmaufnahmen insbesondere die Ermittlung der Perspektiventoleranz bzw. der Anzahl der aufzunehmenden Perspektiven erforderlich. Das ist die vertikale Position bzw. der vertikale Winkel der aufnehmenden 3D Kamera in Bezug zum Streichorchester bei dem noch plausibel wirkende, stereoskopisch funktionierende Filmaufnahmen möglich sind.



Abbildung 88: 3D Bild des in Greenscreenbox spielenden Streichquartetts; das nur mit 3D Brille stereoskopisch zu betrachtende 3D Bild wurde zur Kontrolle vom 3D Camcorder Panasonic AG 3D A1 an einen 3D fähigen Monitor weitergeleitet

5.1 Ermittlung der Perspektiven

Da die Zuschauerperspektive auf das Streichquartett in jedem Konzertsaal unterschiedlich ist, musste auch das Streichquartett in unterschiedlichen Perspektiven aufgenommen werden. Die wesentlichen Gründe dafür waren, dass die Streichquartett-Perspektiven mit der jeweiligen bereits aufgenommenen Panoramahintergrund-Perspektive weitgehend übereinstimmen und Trapezverzerrungen vermieden werden mussten, damit die Plausibilität erhalten blieb. Trapezverzerrungen entstehen bei senkrechter Bildbetrachtung von einer höheren bzw. tieferen Position aus, wodurch die Oberseite scheinbar länger und die Unterseite kürzer bzw. umgekehrt wird (Abbildung 89). Darüber hinaus wird das Bild horizontal gestaucht. Dieser Aufsichtswinkel-Fehler tritt auf, wenn das Streichorchester bei der Videoaufnahme nicht aus dergleichen Kamerahöhe wie die mittels des FABIAN-Systems aufgenommenen Konzertsaalpanoramen erfasst würde, da beide beim Visual Composing (siehe Abschnitt 5.4) miteinander verknüpft werden. Da es aus Zeit- und Kostengründen nicht möglich war für jeden der sechs Konzertsäle eine perspektivische Streichquartett-Aufnahme zu realisieren, sollte die Perspektiventoleranz ermittelt werden, also der Winkelbereich, in dem unterschiedliche Perspektiven noch als plausibel betrachtet werden können. Um die Wirkung des perspektivischen Eindrucks bei verschiedenen Betrachtungswinkeln zu ermitteln, wurden Abschätzungen durchgeführt. Dabei wurde deutlich, dass ab ca. 5° die Plausibilität hinsichtlich der Perspektivwirkung nachlässt und verstärkt Trapezverzerrungen auftreten. Bei einer Probemontage des Streichquartetts mit einem Aufsichtswinkel von -10° (entspricht Betrachtung von oben) gegenüber dem Konzertsaalhintergrund ist ein „Kippen“ des Streichquartetts nach vorn zu erkennen, was bereits zu einem Verlust an Plausibilität führt (Abbildung 90).

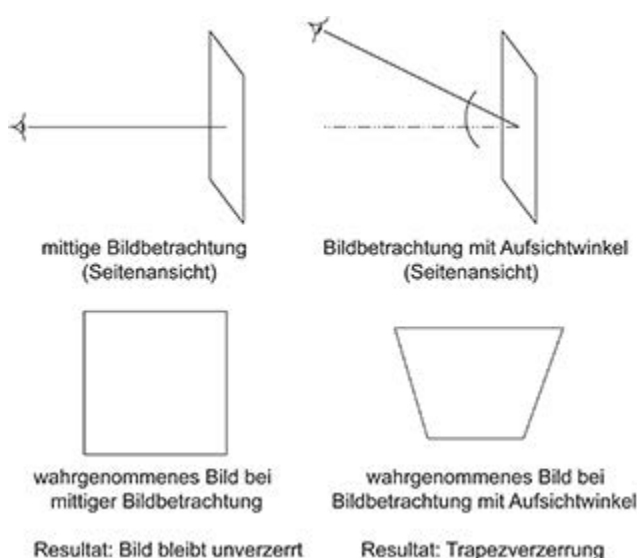


Abbildung 89: Effekt der Trapezverzerrung

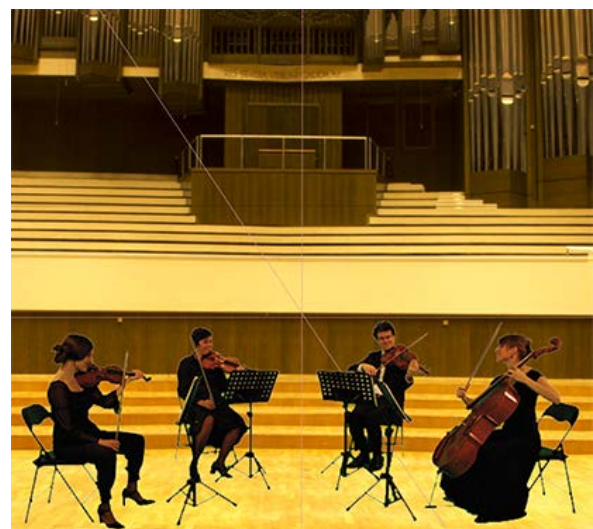


Abbildung 90: Streichquartett hat gegenüber dem Konzertsaalhintergrund einen Aufsichtswinkel von -10° (nach vorn gekippt; entspricht Betrachtung von oben)

Um die verschiedenen Aufsichtswinkel der Konzertsaalpanoramen zu ermitteln, wurde die Geometrie der Panoramen bei der Akquise mit erfasst. Aus den gesammelten Daten konnten dann die Aufsichtswinkel für das Streichquartett mittels trigonometrischer Gleichungen berechnet werden (Anhang A5). Im Folgenden werden die Musiker auf der Bühne als Sender und die aufzeichnende Kamera in Höhe des Publikums als Empfänger bezeichnet. Aus den drei gemessenen Werten für die Senderhöhe **S**, die Empfängerhöhe **E** und den Abstand zwischen Sender und Empfänger **B** konnte der Aufsichtswinkel α für die Elevation¹ berechnet werden:

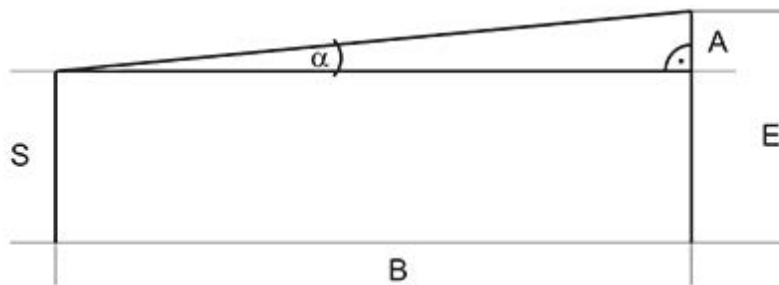


Abbildung 91: Bestimmung der Streichquartett-Aufsichtswinkel mittels Geometrie der Konzertsaalpanoramen

$$\tan \alpha = \frac{A}{B} \quad \rightarrow \quad \alpha = \arctan \frac{A}{B} \quad \rightarrow \quad \alpha = \arctan \frac{S - E}{B}$$

$$\alpha = \arctan \frac{S - E}{B}$$

- α – Aufsichtswinkel
- S** – Senderhöhe
- E** – Empfängerhöhe
- A** – Höhendifferenz zwischen Sender und Empfänger
- B** – Abstand zwischen Sender und Empfänger

Bei der Verwendung der Streichquartett-Mitte als hohes Zentralziel auf der Bühne mit der Senderhöhe **S = 1,11m** wurden für die Konzertsäle und den ursprünglich für die CAVE vorgesehenen Seminarraum folgende Ergebnisse ermittelt:

	GH	JCK	KE	KH	KO	RT	HFT
Konzertsaal	Gewandhaus Leipzig	Jesus-Christus-Kirche Berlin	Kloster Eberbach Basilika	Konzerthaus Berlin Kleiner Saal	Komische Oper Berlin	Renaissance-Theater Berlin	Seminarraum HFT-FT617
E [m] Empfängerhöhe	2,063	1,236	1,236	0,236	0,58	0,754	1,236
B [m] Abstand zwischen Sender und Empfänger	9,863	7,2	15,84	9,888	9,42	9,64	3,6
alpha [°] Aufsichtswinkel	-5,52	-1,0	-0,45	5,1	3,2	2,11	-2,0

Abbildung 92: Tabelle zur Bestimmung der Aufsichtswinkel aus Sender-/Empfängerhöhe und Abstand

¹ Höhenwinkelabweichung

Das Vorzeichen bestimmt, ob der jeweilige Aufsichtswinkel oberhalb oder unterhalb der Nulllinie liegt. Nach maßstabsgerechter grafischer Erfassung dieser Ergebnisse ergaben sich folgende Aufsichtswinkel:

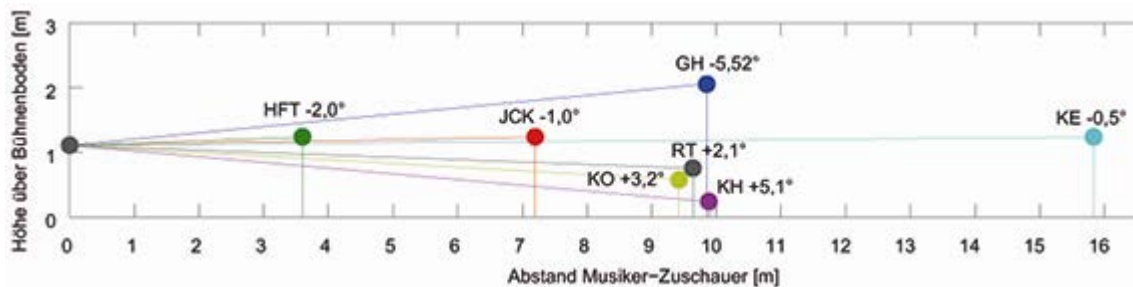


Abbildung 93: maßstabsgerechte grafische Erfassung der Aufsichtswinkel aller Konzertsäle

Aus Abbildung 93 geht hervor, dass sich alle Aufsichtswinkel in einem Bereich zwischen $-5,52^\circ$ und $+5,1^\circ$ bewegen, was einer Differenz von $10,62^\circ$ entspricht. Da wie auf S. 75 beschrieben die Plausibilität bei Aufsichtswinkeln ab ca. 5° deutlich geringer wird, bot sich bei einem vertretbaren Aufwand an Zeit und Kosten die Möglichkeit, drei Perspektiven in drei Empfängerhöhen für alle Konzertsäle zu erfassen, sodass noch eine plausible Wahrnehmung gewährleistet werden kann. Dazu wurden benachbarte Aufsichtswinkelwerte, auch als Elevation bezeichnet, zusammengefasst und deren Mittelwerte berechnet. Als Grundlage für die weitere Berechnung dient wieder die trigonometrische Beziehung aus Abbildung 91. Nachdem die Durchschnittswerte für die Aufsichtswinkel berechnet wurden, kann nun für jede der drei Perspektiven A, B und C die Empfängerhöhe E und letztlich die jeweilige Kamerahöhe für die stereoskopische Videoaufnahme berechnet werden:

$$\tan \alpha = \frac{S - E}{B} \quad \rightarrow \quad S - E = \tan \alpha \cdot B \quad \rightarrow \quad -E = (\tan \alpha \cdot B) - S \quad \rightarrow \quad E = S - (\tan \alpha \cdot B)$$

Für die Berechnung wurden die Konzertsäle je einer von drei Perspektiven zugeordnet:

Perspektive A: Konzerthaus Berlin, Komische Oper Berlin, Renaissance-Theater Berlin

Perspektive B: Jesus-Christus-Kirche Berlin, Kloster Eberbach, Seminarraum HFT-FT617

Perspektive C: Gewandhaus Leipzig

In der folgenden Tabelle sind die Berechnungen für die Aufnahmehöhen der drei Konzertsaal- Perspektiven zusammengefasst:

Perspektive	Gültig für Räume	Mittlere Elevation $\varnothing \alpha$ [°]	Richtung +1=Untersicht -1=Aufsicht	Mittlerer Abstand zwischen Sender und Empfänger $\varnothing B$ [m]	Aufnahmeabstand [m]	Aufnahmehöhe [m]
A	KH, KO, RT	3,60	1,00	9,65	8,15	0,60
B	HFT, JCK, KE	1,20	-1,00	8,88	8,15	1,28
C	GH	5,52	-1,00	9,86	8,15	1,90

Abbildung 94: Tabelle zur Bestimmung der Aufnahmehöhe aus Aufnahmeabstand und mittlerer Elevation bei einer Zentralquelle $S = 1,11\text{m}$

Bei der Berechnung ist zu beachten, dass in die Ermittlung des Tangens innerhalb der obigen Gleichung das in Abbildung 94 in Spalte 4 (Richtung) angegebene Vorzeichen mit einfließt. Wie aus Abbildung 94 ersichtlich ist, weicht der tatsächlich verwendete Aufnahmeabstand von dem mittleren Abstand zwischen Sender und Empfänger ab. Dies war letztlich den räumlichen Gegebenheiten geschuldet, die keinen größeren Abstand als 8,15m zuließen. Nach Einsetzen dieses Wertes für **B** in obige Gleichung konnten drei verschiedene Aufnahmhöhen berechnet und realisiert werden. Als Vorteil erwies sich dabei, dass die Kamera mit dem gleichen Aufnahmeabstand bei der Einstellung der unterschiedlichen Perspektiven nicht umgesetzt werden musste, wodurch Zeit gespart werden konnte. Durch die einfache Höhenverstellung des Kamerastativs entsprechend den berechneten Aufnahmhöhen (Abbildung 94) konnten somit technisch korrekte Einstellungen für die sieben Streichquartett-Perspektiven in Bezug zu den sechs Konzertsälen und dem Seminarraum mit einer tolerierbaren Perspektivendifferenz von maximal 5° für die spätere Montage mit den Konzertsaalpanoramen realisiert werden.

5.2 Vorbereitung

Nachdem im Vorfeld der stereoskopischen Videoproduktion die im letzten Abschnitt erläuterten Streichquartett-Aufnahmeperspektiven ermittelt wurden, mussten ein Aufnahme-studio und eine 3D Videoproduktionsfirma gefunden werden. Eine Zusammenarbeit kam schließlich mit der Firma „Yourcity3D“ und den „FrogFish Studios“ in Berlin zustande. Des Weiteren wurde eine Materialliste erstellt, die neben technischem Equipment für die Audio-wiedergabe insbesondere Messinstrumente wie Graukarte/Weißkarte für die Farbtemperatur und Lasermeter für die Entfernungsbestimmung enthielt (Anhang A6). Akquirierte Noten-ständer und Stühle mussten entsprechend der in Abbildung 87 dargestellten Geometrie angeordnet werden und sollten im Laufe der Aufnahmen möglichst nicht verändert werden (Abbildung 95). Für eine plausible Wiedergabe von Bild und Ton musste sich das Streichquartett darauf einstellen, synchron zu dem bei der nachhallfreien Audioproduktion aufgenommenen Stück synchron zu spielen. Diese und weitere Fragen zu den Konditionen der Videoproduktion wurden im Vorfeld zwischen dem Projektleiter, der Videoproduktions-firma, den Musikern und der Assistenz besprochen. Ein zentrales Thema war dabei die Wahl eines geeigneten stereoskopischen Camcorders. Als ein qualitativ hochwertiger und dennoch kompakter und leicht handhabbarer stereoskopischer Camcorder sollte der Panasonic AG-3DA1 zum Einsatz kommen (Abbildung 96). Dieses voll integrierte Full HD 3D Gerät mit einer Auflösung von 1920x1080 Pixeln und zwei Objektiven ermöglicht eine Aufzeichnung mit einer Vollbildwiederholfrequenz von 30 Bildern pro Sekunde auf zwei SDHC Speicherkarten im AVCHD Format mit der Datenkompression MPEG-4 AVC/H.264

¹ Advanced Video Coding High Definition – Videoformat mit hoher Auflösung

und gleichzeitig eine einfache Bedienung der stereoskopischen Parameter, insbesondere die Justierung des Konvergenzpunktes (Vgl. Abschnitt 3.2.3.1, S. 41). Dabei kommen die eher kleinen 1/4,1" großen 3 Farb- CMOS-Sensoren (Vgl. Abschnitt 3.2.1, S. 29 ff) mit einer Auflösung von 2.07 Megapixel je Farbe und je Auge zum Einsatz. Der Brennweitenbereich des 5,6 fachen Zoomobjektivs liegt zwischen $f=4,2\text{mm} - 23,5\text{mm}$ bzw. auf das 35mm Kleinbildformat umgerechnet zwischen $47,1\text{mm} - 264\text{mm}$ [88, S. 92-93].



Abbildung 95: Greenscreen-Studio mit Plätzen und Notenständern für Streichquartett und Target-Karte für Fokussierung und korrekte Positionierung in der Postproduktion



Abbildung 96: 3D Camcorder Panasonic AG-3DA1 zur stereoskopischen Aufnahme des Streichquartetts

5.3 Durchführung

Entsprechend der Planung wurde die stereoskopische Greenbox-Videoaufnahme des Streichquartetts am 11.12.2011 von 14.00-19.00 in den „FrogFish Studios“ (Manager: Fabian Bollig) Fredersdorfer Straße 10, 10243 Berlin unter der Leitung von Dr. Hans-Joachim Maempel (TU Berlin, SEACEN-Projekt) durchgeführt. Aus Gründen der Plausibilität und Synchronität wurde wie bei der nachhallfreien Audioproduktion das Berlin-Budapest Quartett bestehend aus: Dea Szücs (1. Violine -VI1), Éva Csermák (2. Violine -VI2), Itamar Ringel (Bratsche -VIa) und Ditta Rohmann (Cello -Vc) engagiert. Für die stereoskopische Videoaufnahme war die Produktionsfirma „Yourcity3D“ unter der Regie von Shamir Ali-Khan und der Mitwirkung von Kameramann Alexander Haßkerl und Frau Annika Natus verantwortlich. Von der TU Berlin waren Frau Mina Fallahi und der Autor als Assistenz vor Ort.

Bevor das Streichquartett aufgenommen werden konnte, musste die Greenbox aufgebaut und das technische Equipment wie Camcorder und Lichttechnik eingerichtet werden. Für eine geeignete Beleuchtung der Szenerie kamen zahlreiche Strahler, Fluter etc. mit einer Leistung von insgesamt 24 kW zum Einsatz (Abbildung 99, Anhang A7). Um eine Aufnahme zu ermöglichen, die mit den Konzertsaalpanoramen korrespondiert, war die genaue Kalibrierung aller Parameter notwendig. Dazu gehörten insbesondere die optischen Parameter des stereoskopischen Camcorders wie: Brennweite, Belichtungszeit, Blende,

Fokus, ISO- bzw. Gain-Wert, Farbtemperatur, Bildqualität, Bildformat, Stereobasis und Konvergenz.

Die *Brennweite*, die am Camcorder eingestellt werden musste, konnte mit Hilfe der Abbildungsgleichung berechnet werden:

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{b} \rightarrow f = \frac{g \cdot b}{g + b} \quad \text{I}$$

f = Brennweite - Abstand zwischen Brennpunkt und Linsenmittelpunkt
b = Bildweite - Abstand zwischen Linse und Bildsensor
g = Gegenstandsweite - Abstand zwischen Linse und Gegenstand

Mit der Beziehung des Strahlensatzes entsprechend Abbildung 39 (S. 31):

$$\frac{b}{B} = \frac{g}{G} \rightarrow b = \frac{g \cdot B}{G} \quad \text{II}$$

B = Höhe des Camcorder-Bildsensors
G = Höhe des aufzunehmenden Gegenstandes, also des Streichquartetts

erhält man durch Einsetzen von **II** in **I** und der damit verbundenen Eliminierung von **b**:

$$f = \frac{B \cdot g}{B + G} \quad \text{III}$$

Bei der sitzenden Position des Streichquartetts ergibt sich eine durchschnittliche Gegenstandshöhe **G** von 1,55 m. Der ¼" große Bildsensor hat eine Sensor-Bildhöhe **B** von 2,4 mm. Die Gegenstandsweite **g**, also der Abstand zwischen Gegenstand (Streichquartett) und optischem Zentrum des Camcorders beträgt entsprechend Abbildung 94 (S.77) 8,15 m. Damit wird Gleichung **III** zu:

$$f = \frac{2,4 \text{ mm} \cdot 8,15 \text{ m}}{2,4 \text{ mm} + 1,55 \text{ m}} = \frac{2,4 \text{ mm} \cdot 81500 \text{ mm}}{2,4 \text{ mm} + 15500 \text{ mm}} = \frac{195600 \text{ mm}}{15502,4 \text{ mm}} = 12,62 \text{ mm}$$

Der Brennweiten-Wert **f = 12,6 mm** kam für die mittlere und hohe Perspektive B und C mit einem Zoomfaktor von 33 (B) bzw. 35 (C) zum Einsatz. Um gleiche Größenverhältnisse für eine Brennweite von 4,2 mm bei Perspektive A zu erhalten, wurde ein Zoomfaktor von 11 verwendet, da eine Verkleinerung der Brennweite eine um den gleichen Faktor (in diesem Fall um den Faktor 3) proportionale Verringerung des Zoomfaktors nach sich zieht.

Die Erfassung der *Belichtungszeit*, mit deren Hilfe auch die *Blende* und die Lichtmenge bzw. der Lichtwert bestimmt wird, kann mittels einer Lichtmessung oder einer Objektmessung erfolgen. Während bei der Lichtmessung vom Motiv ausgehend das direkt einfallende Licht gemessen wird, kommt es bei der in diesem Fall angewandten Objektmessung zur Erfassung des vom Motiv reflektierten Lichtes bzw. dessen Leuchtdichte. Durch Positionierung einer Graukarte an die Stelle des Motivs kann eine Ersatzmessung durchgeführt werden. Die Graukarte als Objekt bekannter Helligkeit, reflektiert 18 % des auf sie einfallenden Lichtes, was im Durchschnitt der Lichtmenge bei einer direkten Lichtmessung entspricht. Mittels dieser Graukarten-Ersatzmessung ist eine ausreichend

genaue, der direkten Lichtmessung nahe kommende Erfassung der Lichtwerte möglich [89]. Die Belichtungszeit (exposure time) wurde auf 1/50s und die Verschlussgeschwindigkeit (shutter speed) auf 59,9 Hz eingestellt.

Die *Fokussierung* sollte mittels Target-Karte so erfolgen, dass das Streichquartett scharf abgebildet wird. Die Empfindlichkeit des Bildsensors wird durch den *Gain-Wert* bestimmt, der auf den kleinstmöglichen Wert von 0dB eingestellt wurde. Bei Erhöhung des Gain-Wertes steigt die Empfindlichkeit, gleichzeitig jedoch auch das Bildrauschen.

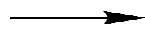
Um den *Weißabgleich*, also die Anpassung der Farbtemperatur des Streichquartetts an die Farbtemperatur der Stereopanoramen zu gewährleisten (Vgl. Abschnitt 3.1.3, S. 21-22), wurden Fotos einer an die Stelle der Musiker positionierten Weißkarte aufgenommen. Diese ermöglichten in der Postproduktion den weiteren Farbabgleich.

Die *Bildqualität* wird im Wesentlichen durch die Bildauflösung bestimmt. Diese entspricht der gegenwärtig im Kamerabereich höchsten Ansprüchen genügenden Full HD Auflösung von 1920 x 1080 Pixeln = 2.07 Megapixeln. Die Aufnahmen wurden im AVCHD PH mode, 1080/30p*2 Bildformat mit der Datenkompression MPEG-4 AVC/H.264 auf SDHC Karten gespeichert. Dabei wurde für jedes Auge in Full HD auf je eine Speicherkarte aufgenommen [88, S. 92-93].

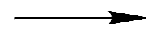
Die *Stereobasis* ergibt sich aus dem Abstand der optischen Achsen der beiden Objektive voneinander. Dieser betrug ca. 6 cm und orientierte sich damit an dem durchschnittlichen menschlichen Augenabstand von 6,5 cm.

Da der Abstand zwischen Camcorder und Streichquartett mit 8,15 m deutlich über 2 m lag, wurde die *Konvergenz* auf parallel und damit ein Konvergenzwinkel von 0° eingestellt (Vgl. Abschnitt 2.3, S. 11-12; Abbildung 14). Bei dieser *parallelen Ausrichtung* der beiden Kameraobjektive liegt die so genannte Nullparallaxe im Unendlichen (Abbildung 97). Das ist die Ebene, in der sich die Strahlen der beiden optischen Achsen der Objektive schneiden und auch keine parallaktische Verschiebung aufgrund der Querdissipation auftritt. Der Vorteil der parallelen Ausrichtung ist, dass der Abstand der stereoskopischen Szenerie von der Nullparallaxe in der Postproduktion einstellbar ist. Bourke zeigt dazu eine Methode, die speziell für stereoskopische Panoramaaufnahmen eingesetzt wird [90, S. 150-152]. Danach werden die beiden Stereo-Teilbilder, wie in Abbildung 98 dargestellt, mittels einer Drehung der virtuellen Objektive um die Drehachse in der Postproduktion gegeneinander verschoben (rot), bis der gewünschte Abstand zur Nullparallaxe erreicht wird. Diese Kalibrierung erfolgte so, dass die Stereo-Szenerie bei der späteren Projektion in der CAVE *hinter* der Leinwand wahrgenommen wird. Zur genauen Berechnung kam ein trigonometrischer Ansatz des Sinus' bzw. des Arcussinus' zur Anwendung. Mit Hilfe des halben Objektivabstandes r und der Nullparallaxe f_0 konnte der Winkel zwischen den Kameras Φ folgendermaßen berechnet werden:

$$\sin \frac{\Phi}{2} = \frac{r}{f_o}$$



$$\frac{\Phi}{2} = \arcsin \frac{r}{f_o}$$



$$\Phi = 2 \arcsin \frac{r}{f_o}$$

[90, S. 150-152]

Bei der *konvergenten Ausrichtung* der beiden Objektive wäre die Entfernung zur Nullparallaxe schon bei der Aufnahme festgelegt worden, wodurch eine postproduktive Anpassung nicht mehr möglich gewesen wäre. Eine Berechnung wäre dann über den Arcustangens möglich, da f_o hier von der Hypotenuse zur Ankathete wird (Abbildung 98, grünes Dreieck).

$$\tan \frac{\Phi}{2} = \frac{r}{f_o}$$



$$\frac{\Phi}{2} = \arctan \frac{r}{f_o}$$



$$\Phi = 2 \arctan \frac{r}{f_o}$$

[13, S. 32]



Abbildung 97: Frontansicht des verwendeten Panasonic AG-3DA1 Camcorders mit parallel zueinander ausgerichteten Objektiven

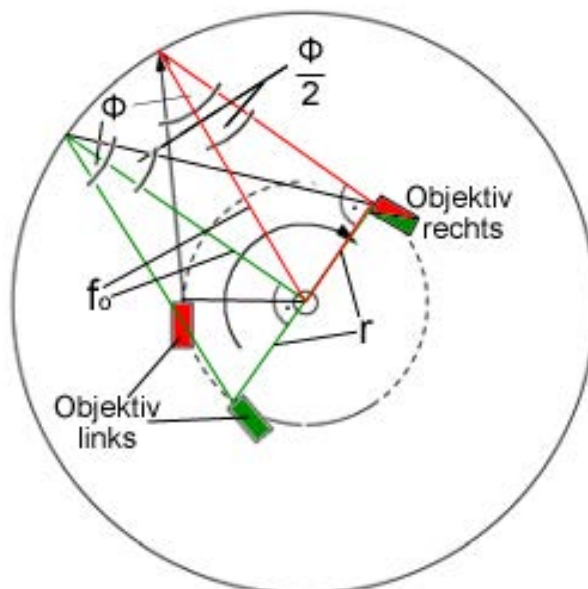


Abbildung 98: Bestimmung des Parallaxenwinkels und damit des Abstandes zur Nullparallaxe durch Verschiebung der beiden Stereoteilbilder in Form virtueller Objektive; rot: parallele Objektivausrichtung grün: konvergente Objektivausrichtung [90, Vgl. Abb. S. 152]

Da die Video- und Audioaufnahmen wegen der unterschiedlichen Aufnahmebedingungen (insbesondere Greenscreen-Hintergrund, reflexionsarmer Raum) getrennt voneinander erfolgen mussten, war die *Synchronisierung* von Video und Audio eine wichtige zu lösende Aufgabe. Die Audioaufnahmen zu bereits vorhandenen Videoaufnahmen zu synchronisieren wäre nicht praktikabel gewesen. Dabei hätte bei der Audioaufnahme im reflexionsarmen Raum jedem Musiker ein Monitor mit den vorhandenen Videoaufnahmen zur Ansicht des eigenen Spiels als Vorlage zum Synchronspiel bereitgestellt werden müssen. Abgesehen von den technischen Herausforderungen, hätten sich die Musiker stark auf die Videoaufnahmen konzentrieren müssen, wodurch möglicherweise die emotionale Ausdrucksfähigkeit

des Spielens eingeschränkt worden wäre. Aus diesen Gründen erfolgte die Synchronisierung des zu filmenden Streichquartetts zu den für die CAVE bereits vorliegenden im nachhallfreien Raum erfassten Audioaufnahmen. Die Herausforderung für die Musiker bestand darin, zu den vom CD-Player über zwei Beschallungslautsprecher vom Typ QSC wiedergegebenen hallfreien Audioaufnahmen des 2. Satzes für Streichquartett von Claude Debussy synchrone, insbesondere spieltechnisch authentische Bewegungen über die gesamte Dauer des Stückes auszuführen (Abbildung 100).



Abbildung 99: Beleuchtung bestehend aus Strahlern, Flutern etc. ; Leistung: 24 kW



Abbildung 100: Berlin-Budapest Quartett spielt Stück von Debussy vor Greenscreen-Wand

Nachdem sich die Musiker entsprechend den Positionen der einzelnen Streichinstrumente platziert hatten (Vgl. S.72 Abb. 87) und die Beleuchtung bzw. der Camcorder kalibriert waren, konnten die Aufnahmen beginnen. Da für die Wiedergabe in der CAVE keine Bildschnitte bzw. –sprünge innerhalb einer Darbietung erlaubt sind, musste der ca. vier Minuten lange Streichquartett-Satz in einem Stück durchgespielt und aufgenommen werden. Aus diesem Grunde wurden so viele Takes erfasst, bis unter der künstlerischen Leitung von Dr. Maempel ein finaler spieltechnisch und aufnahmetechnisch nahezu fehlerfreier Take zustande kam.

Nachdem die exakte horizontale Ausrichtung des Camcorders mit Hilfe einer Wasserwaage erfolgt war, sollte das Streichquartett (Vgl. Abschnitt 5.1, S. 75) aus drei Perspektiven aufgenommen werden. Es musste also für jede der drei Perspektiven jeweils ein finaler Take erfasst werden. Zunächst wurde von Kameramann Alexander Haßkerl die mittlere Perspektive B in Höhe von 1,28 m eingestellt und entsprechend den Regieanweisungen von Dr. Maempel und Shamir Ali-Khan aufgenommen. Die Perspektive C in Höhe von 1,90 m und die Perspektive A in Höhe von 0,60 m konnten nur mit Hilfe eines flexiblen vom Autor zur Verfügung gestellten Stativs eingerichtet und erfasst werden (Vgl. S. 77, Abbildung 94; Abbildung 101). Wegen der örtlichen Gegebenheiten und um die Kamera nicht umsetzen zu müssen, wurden alle drei Perspektiven im gleichen Abstand von 8,15 m aufgenommen. So konnte durch einfache Höhenverstellung des Stativs schnell und unkompliziert die jeweilige Perspektive eingestellt werden.



Abbildung 101: Aufnahme der drei Perspektiven; links: Perspektive B - 1,28 m; Mitte: Perspektive C - 1,90 m, rechts: Perspektive A - 0,60 m

Während der Aufnahmen musste der Camcorder so fixiert werden, dass insbesondere Verwackelungen, Bewegungen der Kamera und das Zoomen vermieden wurden.

Zur permanenten Bildkontrolle, insbesondere auch der Stereosituation, kam ein professioneller stereoskopischer Monitor von Zalman, der an dem Camcorder angeschlossen war, zum Einsatz.

Nachdem die in einem Stück aufgenommenen finalen Takes aller drei Perspektiven auf zwei, jeweils für das linke bzw. rechte Auge vorgesehenen SDHC Memory Cards gespeichert worden waren, mussten die Aufnahmen in der Postproduktion so angepasst werden, dass sie in der CAVE eine plausible Wahrnehmung beim Zuschauer hervorrufen. Dieses so genannte *Visual Composing* soll im folgenden Abschnitt näher erläutert werden.

5.4 Postproduktion und Visual Composing

Die stereoskopischen Streichquartett-Greenbox-Aufnahmen sollten in der Postproduktion mittels *Visual Composing* mit den bereits durch das Messsystem FABIAN (Vgl. Abschnitt 3.4.1.2, S. 59) akquirierten sechs stereoskopischen Hintergrundpanoramen (Abbildung 102) verbunden und angepasst werden. Die wichtigsten Kriterien für die Montage waren:

- Erzeugung eines plausiblen Eindrucks
- Vermeidung von sichtbaren Montage-Artefakten
- Einhaltung der vorgegebenen Positionen der Künstler

Um einen plausiblen Eindruck zu erzeugen, mussten insbesondere Parameter wie

Skalierung, Perspektive, Bild- bzw. Projektionswinkel (FOV¹), Farbtemperatur und Helligkeit berücksichtigt und mit den Hintergrundpanoramen in Einklang gebracht werden. Nach dem Übereinbringen mittels berechneter und kalibrierter technischer Parameter erfolgte die Feinmontage letztlich nach ästhetischem Empfinden.

Raumgröße

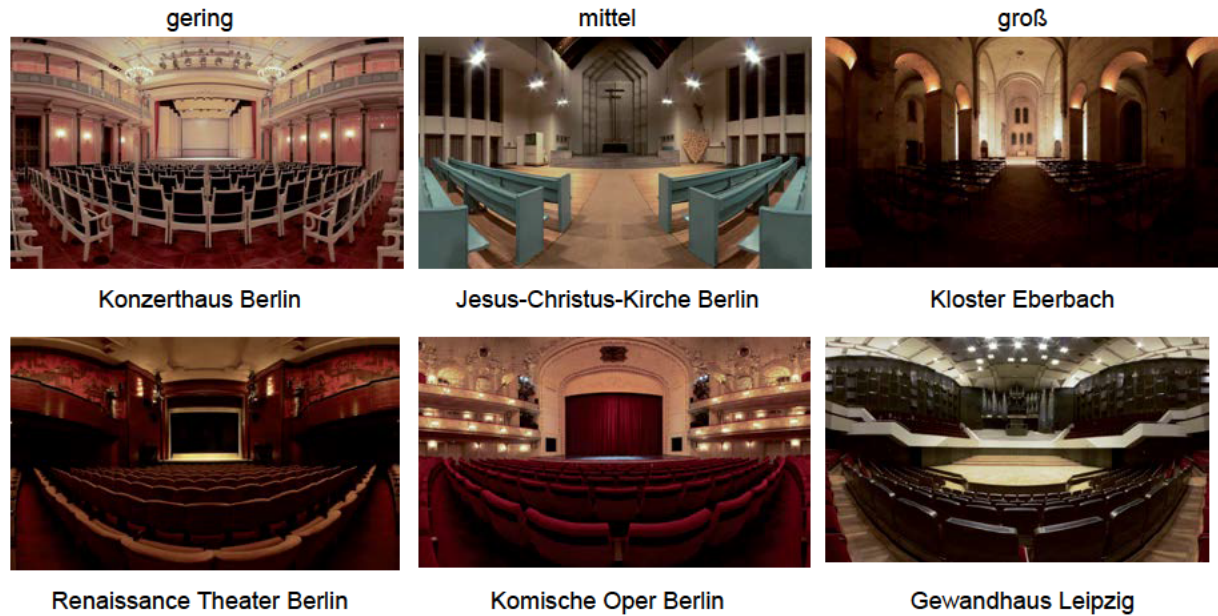


Abbildung 102: Bilder der sechs Hintergrundpanoramen in equirektangularer Projektionsgeometrie; Klassifizierung in für die CAVE wichtige Parameter der Raumwahrnehmung: obere Reihe: geringer mittlerer Schallabsorptionsgrad, untere Reihe: hoher mittlerer Schallabsorptionsgrad; Raumgröße nimmt von links nach rechts zu [76, Vgl. Abb. S. 295]

Die Erzeugung plausibler stereoskopischer Panorama-Projektionen in der CAVE konnte nur durch die konsequente Anwendung digitaler Bildverarbeitungsroutinen erreicht werden. Dazu wurde eine Bildverarbeitungsstrecke konzipiert, die Konvertierungen ohne Qualitätsverlust und schließlich eine unverzerrte Abbildung der stereoskopischen Teilbilder über eine stereoskopische Projektion auf einer zylindrischen Leinwand ermöglichte (Abbildung 103).

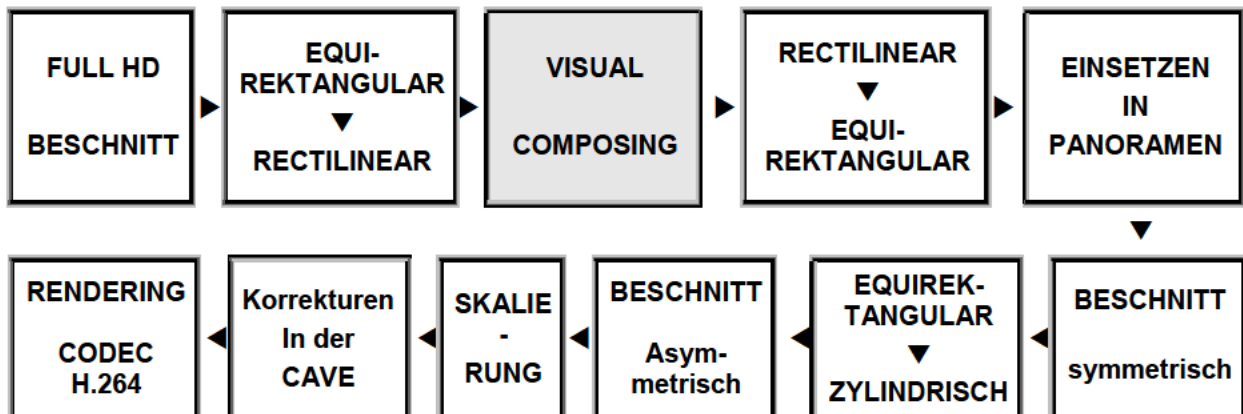


Abbildung 103: digitale Bildverarbeitungsstrecke, die die Erzeugung stereoskopischer Streichquartett-Konzertsaal-Panoramen für die CAVE ermöglicht

¹ FOV – Field of View

Diese Bildverarbeitungsstrecke in der das Visual Composing integriert ist, soll im Folgenden näher erläutert werden:

FULL HD BESCHNITT

Damit die im HD-Format vorliegenden Streichquartett-Aufnahmen mit den $180^\circ \times 180^\circ$ umfassenden und 5400×5400 Pixel großen Hintergrundpanoramen kombiniert werden konnten, wurden die Panoramen in PHOTOSHOP zentriert auf das 1920×1080 Pixel große HD-Format beschnitten.

EQUIREKTANGULAR ► RECTILINEAR

Die Verbindung der Streichquartett-Aufnahmen mit den HD-Panoramenausschnitten mittels Greenscreen-Technik war nur mit gleicher Projektionsgeometrie möglich. Da das Streichquartett mit gnomonischen¹ Objektiven aufgenommen wurde, woraus eine rectilineare² Projektionsgeometrie resultierte, war es notwendig, die in equirektangulärer³ Projektionsgeometrie vorliegenden HD-Panoramenausschnitte (Abbildung 104) in ein rectilineares Format zu wandeln. Dies geschah mit Hilfe der Panorama-Software HUGIN. Nach dem Importieren der HD-Panoramenausschnitte in HUGIN konnten die horizontalen und vertikalen Bildwinkel und die Zielprojektion „rectilinear“ eingegeben werden. Aus diesen Angaben erzeugte das Programm rectilinear gewandelte HD-Panoramenausschnitte (Abbildung 105).



Abbildung 104: Gewandhaus Leipzig; HD-Panoramenausschnitt mit equirektangulärer Projektionsgeometrie



Abbildung 105: Gewandhaus Leipzig; HD-Panoramenausschnitt mit rectilinear gewandelter Projektionsgeometrie

VISUAL COMPOSING

Die Montage der Streichquartett-Aufnahmen mit den HD-Panoramenausschnitten erfolgte durch die Berliner Video-Produktionsfirma Yourcity 3D. Das Visual Composing beinhaltete neben dem Freistellen, Positionieren und Skalieren der Musiker innerhalb der Konzertsaal-Panoramen auch Farb- und Helligkeitsanpassungen des Streichquartetts an die Konzertsäle.

¹ Projektionsart, die verzerrungsfrei, also ohne Verzerrungen abbildet

² geradlinig; verzerrungsfreie Abbildung

³ Projektionsart, bei der Verzerrungen entstehen; ermöglicht u.a. die Abbildung eines ca. 180° großen horizontalen Bildwinkels

Mit Hilfe des Videobearbeitungsprogramms *After Effects* von Adobe und dem Plug in *Keylight* wurden zunächst entsprechend der Greenscreen-Technik die grünen Pixel innerhalb eines festgelegten Toleranzbereiches in transparente Pixel gewandelt. Diesem *Keying* genannten Freistellen des Streichquartetts folgte das Übereinanderlegen von Streichquartett-Videospuren im Vordergrund und HD-Panoramenausschnitten im Hintergrund. Das "Durchscheinen" des jeweiligen Panoramahintergrundes ermöglichte (neben dem mittig aufgestellten Target) die genaue zentrierte Positionierung des Streichquartetts innerhalb des nun sichtbaren Konzertsaales. Auch die drei unterschiedlichen Streichquartett-Perspektiven A, B und C (Vgl. Abschnitt 5.1, S. 77) konnten nun mit den jeweils zugeordneten Konzertsälen direkt verglichen werden. Da jeder der sechs Konzertsaalhintergründe in seiner Perspektive von der jeweils zugeordneten Streichquartett-Perspektive abwich (aufgrund der Durchschnittsbildung der drei Streichquartett-Perspektiven), traten Trapezverzerrungen auf. Diese geringen Abweichungen sollten durch Trapezskalierungen der oberen bzw. unteren Bildbreite des Streichquartetts korrigiert werden (Abbildung 106).

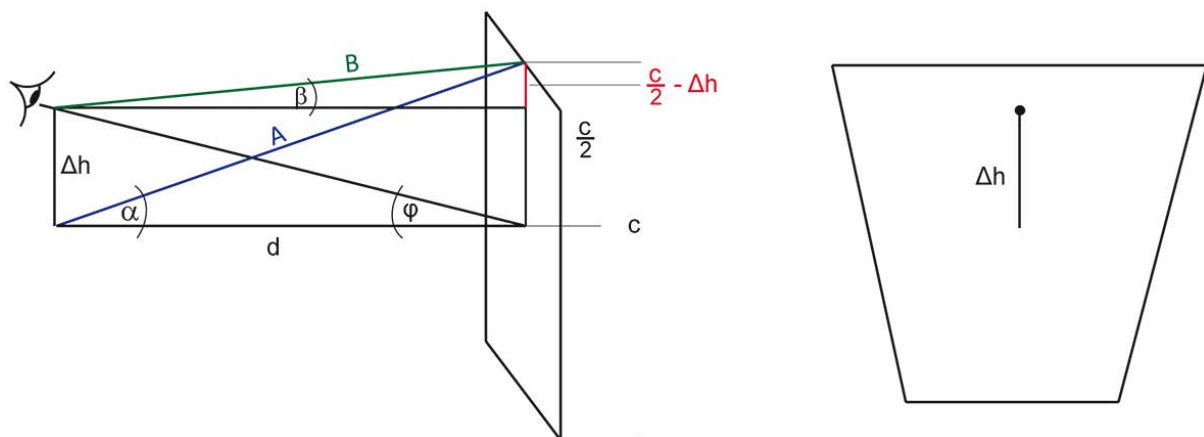


Abbildung 106: Bildbetrachtung mit Aufsichtswinkel und daraus resultierender Trapezverzerrung
links: Seitenansicht; rechts: Ansicht von vorn

Größen:

c = Bildhöhe

d = waagerechter Abstand des Betrachters von der Bildfläche

φ = Aufsichtswinkel; durch Betrachtungshöhenfehler **Δh** verursachter Fehlerwinkel, der zu Trapezverzerrungen führt (Bezug: Waagerechte)

Entsprechend Abbildung 106 wurden mit Hilfe der folgenden Gleichungen Skalierungsfaktoren und Trapezkorrekturwerte für jede Streichquartett-Perspektive in Bezug zu ihrer jeweiligen Konzertsaalhintergrund-Perspektive ermittelt.

Mit

$$\tan \alpha = \frac{c}{2d} \quad \blacktriangleright \quad \alpha = \arctan \frac{c}{2d} \quad \text{und} \quad \tan \varphi = \frac{\Delta h}{d} \quad \text{und} \quad \tan \beta = \frac{c/2 - \Delta h}{d}$$

$$\begin{array}{ccc} \blacktriangledown & & \blacktriangledown \\ \tan \alpha - \tan \varphi = \frac{c}{2d} - \tan \varphi = \frac{c}{2d} - \frac{\Delta h}{d} = \frac{c/2 - \Delta h}{d} & \blacktriangleright & \beta = \arctan \frac{c/2 - \Delta h}{d} = \arctan \left(\frac{c}{2d} - \tan \varphi \right) \end{array}$$

und dem Ansatz:

$$\frac{K_{\text{oben}}}{100\%} = \frac{B}{A} \quad \text{mit} \quad \cos \alpha = \frac{d}{A} \quad \blacktriangleright \quad A = \frac{d}{\cos \alpha} \quad \text{und} \quad \cos \beta = \frac{d}{B} \quad \blacktriangleright \quad B = \frac{d}{\cos \beta}$$

ergibt sich der Skalierungsfaktor K_{oben} zu:

$$K_{\text{oben}} = 100\% \cdot \frac{B}{A} = 100\% \cdot \frac{\cos \alpha}{\cos \beta} \quad \blacktriangleright \quad K_{\text{oben}} = 100\% \cdot \frac{\cos(\arctan(c/2d))}{\cos(\arctan(c/2d) - \tan \varphi)}$$

Für den Skalierungsfaktor K_{unten} , der für die Korrektur der Unterseite ermittelt wurde, gilt:

$$K_{\text{unten}} = 100\% \cdot \frac{\cos(\arctan(c/2d))}{\cos(\arctan(c/2d) + \tan \varphi)}$$

Die ermittelten Skalierungsfaktoren und Werte der korrigierten oberen bzw. unteren Bildbreite (in Pixeln) für die verschiedenen Konzertsäle in Zusammenhang mit den verschiedenen Streichquartett-Perspektiven sind im Anhang A9 zusammengefasst. Mit Hilfe dieser Werte konnten alle geometrischen Fehler kompensiert und die Streichquartett-Aufnahmen perspektivisch korrekt mit den HD-Panoramenausschnitten verbunden werden, sodass ein plausibler Eindruck der Szenerie entstehen konnte.

Da die stereoskopischen Konzertsaalpanoramen in Parametern wie Skalierung, Helligkeit, Farbtemperatur, Auflösung etc. bereits justiert worden waren, durften die HD-Panoramenausschnitte nicht verändert werden, da diese in den nächsten Verarbeitungsschritten wieder in die kompletten Hintergrundpanoramen eingefügt werden sollten. Darum bestand die weitere Anpassung im Wesentlichen in der Bearbeitung der Streichquartett-Aufnahmen unter Einbeziehung der Konzertsaalpanoramen als vergleichende Bildvorlagen.

Die Anpassung der *Größenverhältnisse* der Streichquartett-Aufnahmen in Bezug zu den HD-Panoramenausschnitten konnte durch Skalierung unter Einbeziehung der unterschiedlichen Konzertsaal-Abstände zwischen Sender und Empfänger (Abbildung 92, 93) und den Brennweiten für die unterschiedlichen Perspektiven (Vgl. Abschnitt 5.3, S. 80) realisiert werden.

Um die *Licht- bzw. Beleuchtungsverhältnisse* von Streichquartett-Aufnahmen und HD-Panoramenausschnitten in Übereinstimmung zu bringen, wurden zu Vergleichszwecken Weiß- und Graukarten-Bilder der verschiedenen Konzertsäle bzw. des Streichquartetts und eine Helligkeitstabelle (Anhang A8) verwendet.

Die Farbanpassung mittels *Weißabgleich* erfolgte, indem die Streichquartett-Aufnahmen zunächst mit Hilfe eines ebenfalls im Video-Studio aufgenommenen Weißkartenfotos abgeglichen wurden. Anschließend wurden diese weißabgeglichen Streichquartett-Videos mit Hilfe von Weißkartenfotos der Konzertsäle an die Farbtemperatur des jeweiligen Konzertsaaes angepasst. Zur Endjustage wurden diese Streichquartett-Videos mit vollständigen Konzertsaalpanorama-Bildern verglichen, die die Lichtverhältnisse im gesamten halben Konzertsaal zeigten. Damit konnten die Streichquartett-Aufnahmen bei auftretenden Abweichungen in der Farbtemperatur final angepasst werden.

Die Kalibrierung der *Beleuchtung* der Streichquartett-Aufnahmen erfolgte mit Hilfe der Tabelle Anhang A8, deren Werte für Helligkeitsunterschiede, Blende und Belichtungszeit bzw. Motivhelligkeiten während der Panoramenaquise in den Konzertsälen mittels einer fotografierten Graukarte erfasst wurden. Diese Werte dienen hauptsächlich als Orientierung dafür, ob die Differenzen der Motivhelligkeitswerte der Streichquartett-Graukarten den Differenzen der Motivhelligkeitswerte der Tabelle entsprechen. Die absolute Helligkeit wurde letztlich nach Plausibilität und ästhetischem Eindruck kalibriert. Diese Messwerte und Anpassungen beziehen sich hauptsächlich auf die Bühnenbeleuchtung, die zum Teil sehr hell sein kann, währenddessen der übrige Konzertsaal eher dunkel ist.

Nach der Kombination der sechs Panoramabildausschnitte mit den perspektivisch und fototechnisch angepassten Streichquartett-Filmen, wurden diese bearbeiteten Filmsequenzen als einzelne Panoramaausschnitt-Frames¹ gerendert und als TIFF-Bilder im Full HD Format exportiert (Abbildung 107).



Abbildung 107: Montage von rectilinearen Streichquartett-Aufnahmen mit Panoramabildausschnitten als wesentliche Komponente des Visual Composing am Beispiel des Gewandhauses Leipzig



Abbildung 108: Konvertierung der Streichquartett-Panoramaausschnitt-Frames von einer rectilinearen Projektionsgeometrie in ein euklidisches Format am Beispiel des Gewandhauses Leipzig

¹ einzelne Bilder einer Videosequenz, eine Videosequenz enthält 30 Frames pro Sekunde

RECTILINEAR ► EQUIREKTANGULAR

Um die im rectilinearen Format vorliegenden montierten Streichquartett-Panoramaausschnitt-Frames¹ mit den im equirektangularen Format vorliegenden Halbpanoramen zu verbinden, wurden die Streichquartett-Panoramaausschnitt-Frames mit Hilfe einer Stapelverarbeitungsroutine im Programm HUGIN in das equirektangulare Format konvertiert (Abbildung 108).

EINSETZEN IN PANORAMEN

Mit Hilfe des Matlab-Scriptes „Bilder_zusammensetzen.m“ (Anhang A 10) wurden die einzelnen Panoramaausschnitt-Frames jeder Videosequenz zentriert in die 5400 x 5400 Pixel großen und 180° x 180° umfassenden equirektangularen Halbpanoramen eingesetzt (Abbildung 109, 110).

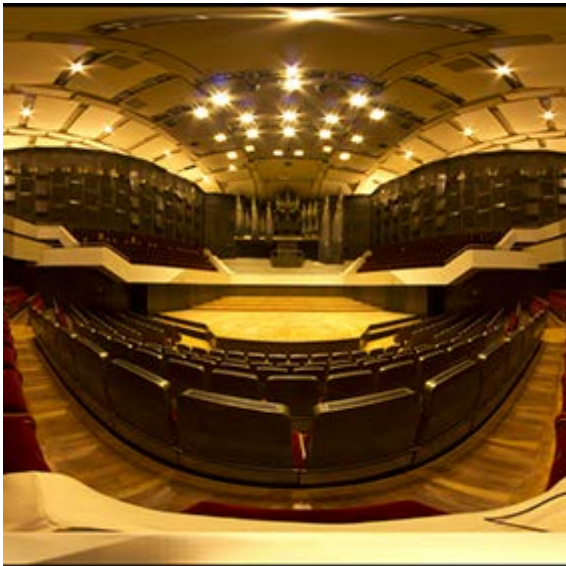


Abbildung 109: Halbpanorama des Gewandhauses Leipzig in equirektangulärer Projektionsgeometrie; FOV: 180° x 180°, Größe: 5400 x 5400 Pixel



Abbildung 110: Halbpanorama des Gewandhauses Leipzig in equirektangulärer Projektionsgeometrie nach dem Einsetzen des 1920 x 1080 Pixel großen Panoramaausschnitt-Frames FOV: 180° x 180°, Größe: 5400 x 5400 Pixel

BESCHNITT symmetrisch

Die Projektion in der CAVE sollte mit fünf HD-Projektoren realisiert werden. Damit eine möglichst hohe vertikale Auflösung und gleichzeitig ein hoch aufgelöstes, horizontales Panorama erreicht werden konnten, wurden die Projektoren nebeneinander im Hochformat angeordnet, sodass die Bildhöhe und -breite vertauscht wurden und sich eine Zielbildgröße von 5400 x 1920 ($5 \times 1080 = 5400$) ergab. Um diese Zielbildgröße zu erreichen, wurden die 5400 x 5400 großen Bilder zunächst beschnitten. Während die Bildbreite von 5400 Pixel erhalten blieb (kein Beschnitt), erfolgte ein symmetrischer Beschnitt auf die vorläufige Zielhöhe von 2336 Pixel, also ein oberer bzw. unterer Beschnitt von jeweils 1532 Pixel.

Diese Zielhöhe ergab sich aus den geometrischen Gegebenheiten der CAVE. Entsprechend Abbildung 111 und 112 konnte bei gegebenen Größen wie Leinwandabmessungen, Abstand des Zuschauers von der Leinwand sowie Länge der Leinwandteilstücke ober- und unterhalb der optischen Achse des Zuschauers zunächst der gesamte FOV des Zuschauers mittels trigonometrischer Zusammenhänge berechnet werden:

$$\tan \alpha = \frac{775}{2500} \quad \blacktriangleright \quad \alpha = \arctan \frac{775}{2500} = 17,22^\circ \quad \text{u} \quad \tan \beta = \frac{2020}{2500} \quad \blacktriangleright \quad \beta = \arctan \frac{2020}{2500} = 38,94^\circ$$

In diesem Schritt wurde zunächst auf den größeren Winkel sowohl oberhalb als auch unterhalb der optischen Achse – also auf $\pm 38,94^\circ$ – symmetrisch beschnitten. Das entspricht bei einer Auflösung von 30 Pixel je Grad einer vertikalen Auflösung von:

$$\pm 38,94^\circ \cong 2 \cdot 38,94^\circ = 77,88^\circ \quad \blacktriangleright \quad 77,88^\circ \cdot 30 \text{ Pixel} = 2336 \text{ Pixel}$$

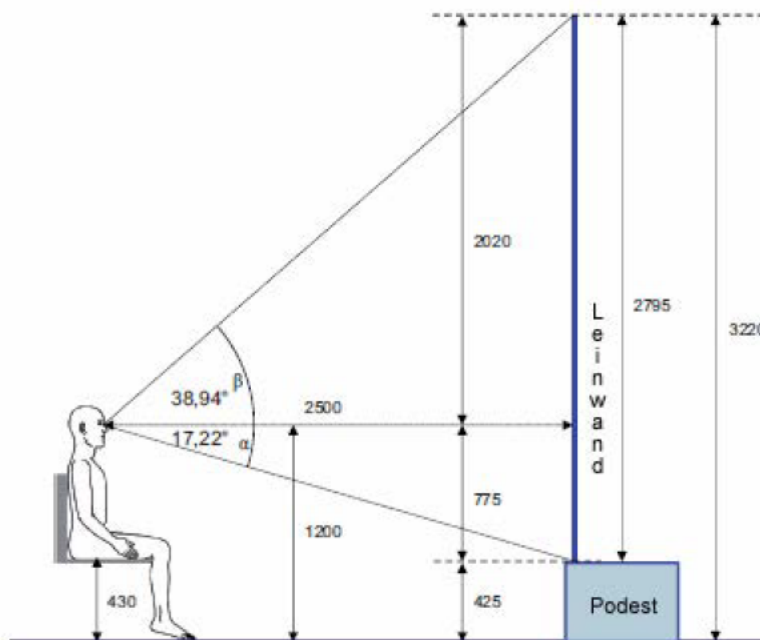


Abbildung 111: CAVE – Seitenansicht [76, Vgl. Abb. S. 304]

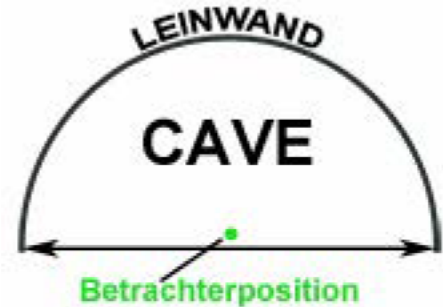


Abbildung 112: CAVE - Draufsicht

EQUIREKTANGULAR \blacktriangleright ZYLINDRISCH

Um eine projektionsgeometrisch korrekte Abbildung auf der in der CAVE verwendeten zylindrischen Leinwand zu erhalten, war die Konvertierung der Streichquartett-Halbpanoramen vom equirektangularen in das zylindrische Format notwendig. Diese Umwandlung der TIFF-Bilder erfolgte mit Hilfe des Programms Hugin (Anhang A11). Als Eingangswerte wurde dabei ein horizontaler FOV von $160,41^\circ$ und ein vertikaler FOV von $77,8^\circ$ gewählt, woraus eine Bildbreite von 4812 Pixel und eine Bildhöhe von 2778 Pixel resultierten. Da mit dieser Konvertierung auch Stauchungen und Streckungen, also letztlich Verzerrungen, einhergehen, änderte sich die vertikale Anzahl von Pixel je Grad. Die Auflösungen ergaben sich zu: horizontal: $4812 \text{ Pixel} / 161,41^\circ = 30 \text{ Pixel je Grad}$

$$\text{vertikal: } 2778 \text{ Pixel} / 77,8^\circ = 36 \text{ Pixel je Grad}$$

BESCHNITT asymmetrisch

Um eine plausible Raumwahrnehmung zu gewährleisten, wurde die Erzeugung eines großen Bildausschnittes angestrebt. Dabei sollten möglichst große Bildwinkel und eine hohe Auflösung erhalten bleiben. Der horizontale Bildwinkel von $161,41^\circ$ weicht mit nur $18,59^\circ$ vom optimalen Bildwinkel von 180° für ein Halbpanorama ab. Für die Wahl des optimalen vertikalen Bildwinkels standen neben dem Streichquartett insbesondere die Konzertsaalwände und –decken im Fokus der Betrachtungen. Während der Panoramen-Akquise in den Konzertsälen war das Aufnahmesystem FABIAN so angeordnet, dass die Objektivposition den Augenpositionen eines sitzenden Konzertbesuchers entsprach. Dabei teilt die horizontale Blickrichtungs-Ebene des Betrachters (Horizontalebene) den vertikalen Bildwinkel in den Winkel β , der oberhalb der Horizontalebene liegt und den Winkel α , der sich unterhalb der Horizontalebene befindet. Um die Konzertsaalwände und den Großteil der Konzertsaaldecken mit abbilden zu können, erschien ein Beschnitt am unteren Rand entlang der Bühnenkante als sinnvoll. Aufgrund der geometrischen Gegebenheiten entsprechend Abbildung 111 war ein vertikaler Beschnitt jedoch unvermeidlich.

Im Verarbeitungsteil *BESCHNITT symmetrisch* wurde bereits der obere Bildwinkel $\beta = +38,94^\circ$ durch Beschnitt festgelegt. Nun sollte der unteren Bildwinkel durch Beschnitt auf seine endgültige Größe gebracht werden. Die Vorgabe für den endgültigen vertikalen Beschnitt war die vertikale HD-Auflösung von 1920 Pixel. Dazu musste die im letzten Konvertierungsschritt entstandene Bildhöhe von 2778 Pixeln um 858 Pixel verringert werden, damit eine vertikale Auflösung von 1920 Pixeln entstand. Von der unteren Bildkante wurden entsprechend 858 Pixel mittels automatisierter Verarbeitungsroutinen in PHOTOSHOP abgeschnitten. Daraus resultierte der untere Bildwinkel $\alpha = 17,22^\circ$. Der gesamte vertikale Bildwinkel –FOV– betrug:

$$\alpha + \beta = 17,22^\circ + 38,94^\circ = 56^\circ$$

Daraus resultierte eine Bildgröße von: 4812 Pixel x 1920 Pixel

horizontal x vertikal

SKALIERUNG

Um den in der CAVE vorliegenden virtuellen Desktop von 5400 Pixel Bildbreite voll auszufüllen, wurde die vorliegende Bildbreite von 4812 Pixel auf 5400 Pixel hoch skaliert. So entstanden 5400 Pixel x 1920 Pixel große Streichquartett-Konzertsaal-Panoramen mit einem Bildwinkel von horizontal 161° x vertikal 56° (Abbildung 113).

Korrekturen in der CAVE

Endgültige Farb- und Helligkeitskorrekturen wurden mit Hilfe der Grau- bzw. Weißkarten-Fotos in der CAVE durchgeführt. Das Hauptkriterium für die Feinjustage war letztlich die perzeptive Plausibilität der Streichquartett-Konzertsaal-Panoramen. Diese Farb- und Helligkeitskorrekturen wurden mittels einer Stapelverarbeitungsroutine in Adobe Bridge mit Camera Raw realisiert.



Abbildung 113: in zylindrisches Bildformat umgewandeltes Streichquartett-Konzertsaal-Panorama des Gewandhauses Leipzig; das an die zylindrische Leinwand der CAVE angepasste Panorama hat eine Bildgröße von 5400 x 1920 Pixel und umfasst einen Bildwinkel von 161° x 56°

RENDERING CODEC H.264

Nachdem die Streichquartett-Panorama-Bilder in Bildgröße und –winkel bzw. in allen foto-technischen Parametern in endgültiger Form vorlagen, mussten sie für jedes Auge zu je einem Film gerendert werden. Dazu musste das nachhallfreie Audiofile jeweils synchron in die verschiedenen Streichquartett-Konzertsaal-Panoramen integriert werden. Als Software-Algorithmus sollte dafür der Videocodec H.264 zum Einsatz kommen. Da H.264 Encoder von kommerziellen Anbietern derzeit keine Filme in der vorliegenden Größe bzw. Auflösung verarbeiten können, wurde die auf dem *QuickTime Animation Format* bzw. auf der Software-Plattform *FFMPEG* [91] basierende, frei verfügbare open source Encodervariante -libx264-, verwendet. Die damit erzeugten relativ kleinen Videodateien lassen sich beispielsweise mit dem VLC Player, einem weit verbreiteten Videoplayer, abspielen.

5.5 Fehlerquellen und Probleme bei stereoskopischen Videoproduktionen

Die Resultate der stereoskopischen Videoproduktion waren von ausreichend hoher Qualität zur Verwendung in der CAVE. Dennoch traten Hemmnisse bzw. Fehler auf, die die Plausibilität, wenn auch nur geringfügig, verminderten. Die zuvor dargestellte Verarbeitungstrecke der stereoskopischen Streichquartett-Panoramen beinhaltete 11 Schritte. Mit jedem Schritt der eine Veränderung von Bildmaterial bzw. deren Parametern nach sich zog, stieg die Fehleranfälligkeit der zu verarbeitenden Daten.

Bedingt durch die Verwendung von drei Aufnahme-Perspektiven, aus denen unterschiedliche Aufnahmen resultierten, unterschieden sich die jeweils in einem Take aufgenommenen Streichquartett-Konzertsaal-Filme geringfügig in der Perspektive und im optischen Spiel der Musiker. Die akustische Performance dagegen war invariant. [Vgl. 76, S. 301]

Eine große Herausforderung für die Musiker bestand (wie in Abschnitt 5.3, S 83 erläutert) darin, ihre optische Darbietung zu dem Playback der Audioaufnahmen synchron in Einklang zu bringen. Da die Aufnahme von Takes solange erfolgte bis die Synchronität der Musiker im Wesentlichen gewährleistet war, traten unvermeidliche, jedoch nur geringe Abweichungen im optischen Spiel der Musiker zwischen den Aufnahmen der drei Perspektiven auf. Diese waren für die Plausibilität jedoch unerheblich und damit nicht relevant.

Beim Einsatz der Greenscreen-Technik zur Freistellung des Streichquartetts war es notwendig, die Entstehung unnatürlicher Effekte wie grünliches Leuchten oder Ausfransen zu verhindern.

Die Beleuchtung musste so eingestellt werden, dass die Musiker nicht überstrahlt wurden. Das wurde durch den Einsatz überwiegend diffuser Strahler ermöglicht (Vgl. S. 83 Abb. 99). Ein pixelgenaues Arbeiten verringerte Fehler, die durch zahlreiche Bild-Transformationsvorgänge beim Konvertieren in verschiedene Projektionsgeometrien bzw. durch die zahlreichen Beschnitt- und Resample-Vorgänge entstehen konnten. Auch durch das Runden von Pixelwerten und darauf folgende Berechnungen konnten Rauschen und Unschärfen entstehen.

Auf besondere Sorgfalt musste bei den Einstellungen der fotometrischen wie auch der stereoskopischen Parameter wie Stereobasis und Konvergenz geachtet werden, da Abweichungen zu unscharfen, farblich verfälschten Bildern bzw. zu einer abweichenden Stereo-Wahrnehmung führen konnten. Insbesondere musste der Wahrnehmungsort von Streichquartett und Panoramen übereinstimmen bzw. angeglichen werden, sodass beide im selben Tiefen-Abstand von der Leinwand wahrgenommen wurden. Des Weiteren gehörte dazu auch die korrekte Anpassung der Größenverhältnisse des Streichquartetts an die Panoramen durch Skalierung und die korrekte Anpassung der Aufsichtswinkel durch Trapezskalierung.

6 Integration der Streichquartett-Panoramen in die CAVE

Aus der im letzten Abschnitt beschriebenen digitalen Bildverarbeitungsstrecke resultierten 5400 x 1920 große stereoskopische Streichquartett-Konzertsaal-Panoramen mit einem Bildwinkel (FOV) von 56° vertikal und 161° horizontal und den dazu synchronisierten binauralen Audiodaten. Für eine plausible Wiedergabe dieser Panoramen und der binauralen Audiaufnahmen in der CAVE musste entsprechendes technisches Equipment akquiriert werden.

Für die visuelle Erfassung sollte ein Projektionssystem zum Einsatz kommen, das die erzeugten stereoskopischen Panoramen auf einer großen Leinwand mit hoher Auflösung und passenden Helligkeits- und Kontrastverhältnissen stereoskopisch plausibel abbildet. Dazu mussten fünf Projektoren gleichen Typs gefunden werden (Vgl. Abschnitt 5.4, S. 90), deren Anschaffung das vorgegebene Budget nicht überschreiten durfte. Nach einem Vergleich der auf dem Markt angebotenen Geräte kamen fünf 3D Full HD LCD-Projektoren der Firma *EPSON* für die CAVE des SEACEN-Projektes zum Einsatz. Das Modell *EH-TW 9000W* ermöglicht eine 3D-Wiedergabe mittels 3LCD-Technologie. Neben einer hohen Lichtstärke bei gleichzeitig guter Farbwiedergabe und umfangreichen 3D Einstellmöglichkeiten gehören eine flexible Aufstellung des Projektors und ein sehr geringes Flimmern aufgrund einer hohen Bildwiederholfrequenz von 480 Hz bzw. 240 Hz pro Auge zu den herausragenden technischen Merkmalen des Gerätes. Dem stehen nur wenige Nachteile wie das Ghosting - das Auftreten von Doppelbildern - gegenüber [92]. Durch die hohe Bildwiederholfrequenz ist dieser Projektor besonders für das Shutterverfahren geeignet, welches eine stereoskopische Wahrnehmung ermöglicht. Durch diese Technik kann mit nur einem Projektor ohne spezielle Anforderungen eine gute Trennung der Stereo-Teilbilder relativ preisgünstig und qualitativ hochwertig realisiert werden. Die Nachteile des Verfahrens wie die aufwendige 120 Hz Technik, die energieabhängige Brillentechnik und der Lichtverlust fielen dagegen nur wenig ins Gewicht (Vgl. Abschnitt 3.2.3.2 S. 44). Als Shutterbrille kam das Modell *Crystal Eyes 4* der Firma *Real D* zum Einsatz. Um die Bilder der fünf Projektoren an den jeweiligen Bildgrenzen korrekt darzustellen, wurden die Bildkanten überlappend projiziert. Die dabei auftretenden Bildfehler wurden mit Hilfe der Software *Pixelwarp* von der Firma *Pixelwix* durch Entzerrung und Kantenglättung korrigiert, sodass ein weitestgehend einheitliches stereoskopisches Konzertsaal-Panoramenbild entstand. Aufgrund der Überlappungen kam es dabei zu Pixelverlusten, sodass das Bild schließlich eine Größe von 4754 Pixel horizontal und 1872 Pixel vertikal bei einer Leinwandhöhe von 3,22 m und einem Radius von 2,50 m erreichte. Diese Werte entsprechen einer so genannten 4K-Projektion mit einer Auflösung von 3840 x 2160 Pixeln, und damit der sowohl horizontal als auch vertikal doppelten Auflösung von Full HD. Gleichzeitig entspricht es etwa dem zweifachen der

Winkelauflösung des menschlichen Auges. Da u.a. das „top and bottom“ 3D Fileformat zum Einsatz kommen sollte, halbierte sich die vertikale Auflösung, da hier die beiden Stereo-Teilbilder übereinander angeordnet sind. Als Steuerzentrale für das Projektionssystem fungiert ein PC mit einem 6-Kern-Prozessor, dem *Intel Core i7 970* mit 3,2 GHz und 1 GB RAM. Der PC besitzt eine schnelle *Powercolor HD 7870 Eyefinity 6* Grafikkarte. Die Wiedergabe der Audio- und Video-Dateien wird mit Hilfe des *VLC-Media-Player* realisiert. Diese CAVE ist dauerhaft im Medialab des Fachbereiches Audiokommunikation der TU Berlin installiert [Vgl. 76, S. 303-304]. Weitere Details zum Aufbau und zur Konfiguration des Projektionssystems können der Masterarbeit von Lisa Esswein am Fachbereich Audiokommunikation der TU Berlin entnommen werden [93].

Für eine plausible Audiowiedergabe in der CAVE ist, wie bereits in Abschnitt 3.4.2, S.61 erläutert, der Einsatz von binauralen Kopfhörern sinnvoll, da diese insbesondere gegenüber Neupositionierungen und anatomischen Unterschieden der Ohr-Morphologie der Testteilnehmer tolerant sind. Außerdem sollten die Kopfhörer neben einem linearen Frequenzgang mit großer Bandbreite das Kriterium der freifeldäquivalente Ohrkanalankopplung (FEC) erfüllen (Vgl. Abschnitt 3.4.2, S.62). Darüber hinaus muss das Kopfhörer-System mit anderen Komponenten wie Mikrofonen, 3D-Brillen bzw. Head-Tracking-Sensoren kompatibel und kombinierbar sein. Das am Fachbereich Audiokommunikation der TU Berlin im Rahmen des SEACEN-Projektes entwickelte BK System erfüllt diese Kriterien. Das System besteht aus einem DSP gesteuerten Leistungsverstärker vom Typ *BKamp* (Abbildung 114) und binauralen Kopfhörern vom Typ *BK 211* (Abbildung 115), die schon für die Montage eines Head Tracking Systems vorbereitet sind. Zu den herausragenden Eigenschaften des BK211 gehört sein geringes Eigenrauschen, das unterhalb der Hörschwelle und ein Klirrfaktor, der über 200 Hz liegt, und eine Übersprechdämpfung, die größer als 23 dB ist. Mit Hilfe des digitalen Signalprozessors (DSP) kann durch Filterung die Kopfhörer-Übertragungsfunktion H_pTF ausgeglichen werden. Die Abweichungen in der linearisierten Kopfhörerübertragungs-

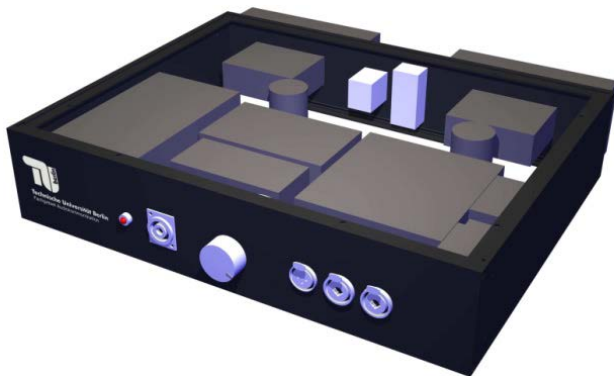


Abbildung 114: DSP gesteuertes Leistungsverstärker vom Typ *BKamp* für den binauralen Kopfhörer *BK 211* [76, Abb. S. 303]



Abbildung 115: an der TU Berlin weiterentwickelter binauraler Kopfhörer vom Typ *BK 211* [76, Abb. S. 303]

funktion sind beim BK211 geringer als bei alternativen, extraauralen, ohrumschließenden Kopfhörern [76, S. 303-305]. Aufgrund dieser Eigenschaften ist das BK System bestens für den Einsatz als akustisches Wiedergabesystem in der CAVE geeignet.

Durch die Kombination aus hochauflösendem, stereoskopischem Panoramafilm und einer binauralen Audiowiedergabe über das BK-Kopfhörersystem wurde eine plausible, virtuelle Umgebung – eine CAVE- erzeugt (Abbildung 116).

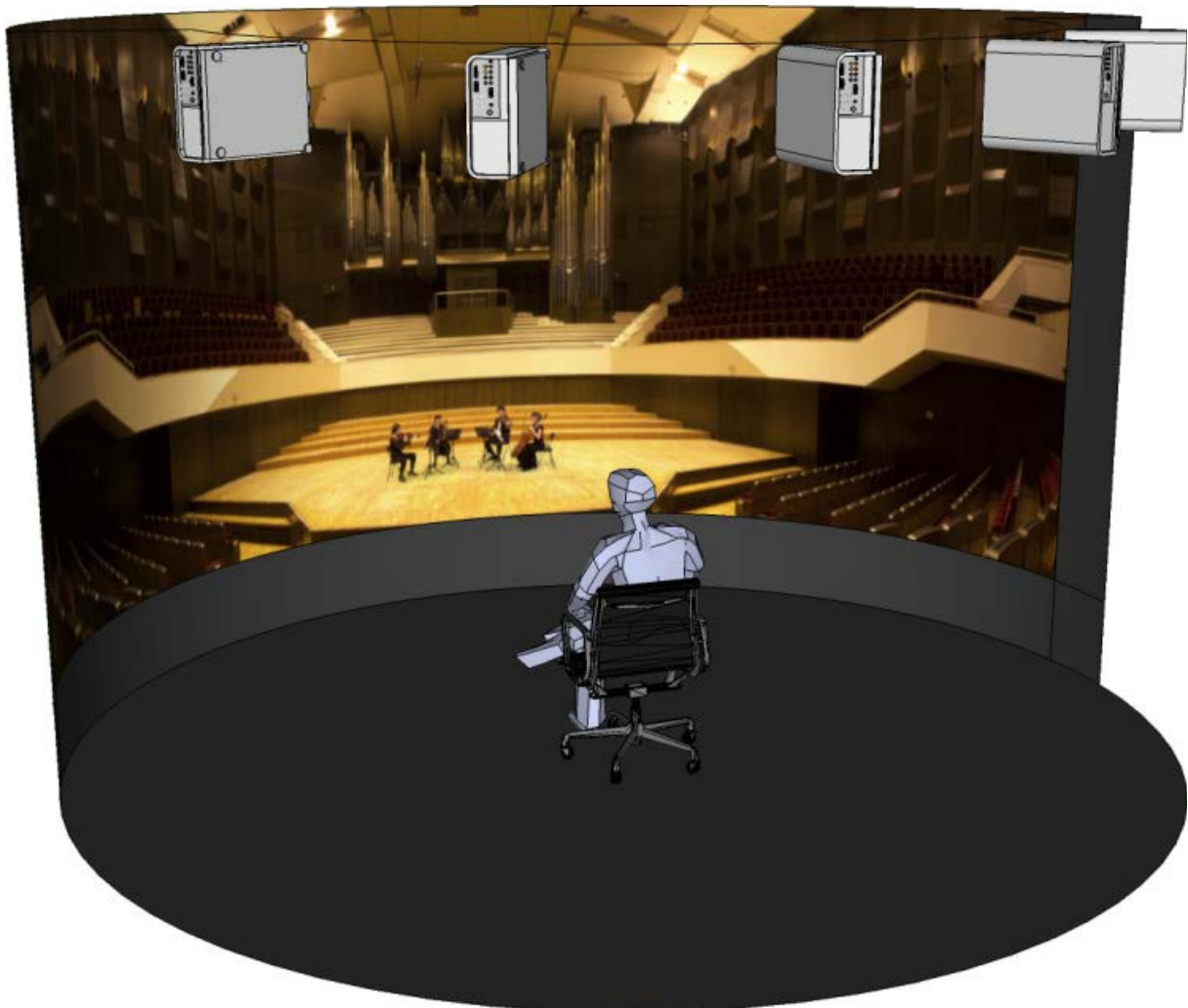


Abbildung 116: schematische Darstellung der CAVE im Medialab des Fachbereiches Audiokommunikation an der TU Berlin [76, Abb. S. 304]

7 Zusammenfassung und Ausblick

Im Fokus des Teil-Projektes P9 der DFG-Forschergruppe Simulation und Evaluation akustischer Umgebungen (SEACEN) stand der Einfluss akustischer und optischer Reize auf die auditive und visuelle Raumwahrnehmung. Zur genauen Untersuchung akustischer und optischer Eigenschaften von Konzertsälen als Stimuli für die visuelle und auditive Raumwahrnehmung und für die Untersuchung von cross modalen (Überschneidungs-) Effekten zwischen visueller und auditiver Raumwahrnehmung sollte eine plausible virtuelle 3D-Umgebung –eine CAVE– geschaffen werden.

Um die Raumwahrnehmung und cross modale Effekte bei Versuchsteilnehmern beobachten zu können, wurde eine spezielle TU-CAVE –das Medialab- erzeugt, das es ermöglicht, optische und akustische Räume unabhängig voneinander variieren zu können. Dazu wurden sechs 161° horizontal und 56° vertikal große, stereoskopische Foto-Panoramen von Konzertsälen sowohl mit deren binauraler Klangcharakteristik als auch mit den stereoskopischen Filmaufnahmen eines Streichquartetts kombiniert.

Im Rahmen der vorliegenden Arbeit sollte dabei nach der Erläuterung der technischen Grundlagen einer virtuellen Umgebung und den psychologischen Grundlagen der visuellen und auditiven Wahrnehmung die nachhallfreie Audioproduktion und die stereoskopische Videoproduktion dargestellt und dokumentiert werden.

Zunächst wurde dazu am Instituts für Technische Akustik an der TU Berlin im reflexionsarmen Raum die nachhallfreie Audioproduktion des *Berlin-Budapest Quartetts* durchgeführt. Um ein gegenseitiges Übersprechen der Instrumente zu verhindern, wurden die Musiker akustisch separiert. Nach der Vorbereitung und dem Entwurf einer Verarbeitungskette erfolgte die nachhallfreie Audioproduktion unter Beachtung von Mikrofonschalleinfallswinkeln und Instrumentenabstrahlwinkeln der einzelnen Streichinstrumente im reflexionsarmen Raum. Dank des technischen Aufbaus und des flexiblen Routings über den Mixer bzw. über die Audioschnittstelle *RME Fireface UFX* konnte der Focus bei den Aufnahmen auf das virtuose, emotionale Spielen der Musiker gelegt werden. Nach der Nachbearbeitung in der Audio-Postproduktion stand das nachhallfreie Audiofile für die Faltung mit den sechs BRIRs der sechs Konzertsäle zur Verfügung, woraus sechs Audiofiles mit der Raumcharakteristik der sechs Konzertsäle resultierten.

Die stereoskopische Videoproduktion, also die Akquise optischer Inhalte, wurde von der Videoproduktionsfirma *Yourcity3D* in den *FrogFish Studios* mittels Greenscreen-Technik realisiert. Neben der Erstellung einer Materialliste und der Vorbereitung der Produktion musste die Perspektiventoleranz für die Aufnahmen ermittelt werden. Nur so konnten unnatürlich wirkende Trapezverzerrungen vermieden werden. Aus den Berechnungen dafür resultierte die Festlegung von drei Aufnahmeperspektiven. Nach dem Aufbau der Greenbox

und der Einrichtung des technischen Equipments wie Camcorder und Lichttechnik erfolgte die genaue Kalibrierung aller optischen Parameter, insbesondere der stereoskopischen Parameter wie Stereobasis und Konvergenz. Eine besondere Herausforderung für die Musiker stellte das synchrone, optische Spielen zu einem vorgegebenen akustischen Playback dar.

Um die so erzeugten Streichquartett-Aufnahmen in eine virtuelle Umgebung einzugliedern, mussten sie zunächst mit den Hintergrundpanoramen der Konzertsäle in der Videopostproduktion, die auch das Visual Composing beinhaltet, zusammenmontiert werden. Dazu wurde eine digitale Bildverarbeitungsstrecke entwickelt, die die gesamte Integration der verschiedenen Streichquartettperspektiven mit den Hintergrundpanoramen in ein zylindrisches Bildformat zur Ansicht in der Medialab-CAVE der TU Berlin ermöglichte. Auch hier erfolgten z.T. Korrekturen der Aufnahmeperspektive durch Trapezskalierungen sowie Korrekturen verschiedener fototechnischer Parameter entsprechend ästhetischen bzw. plausiblen Kriterien. Im Medialab wurde entsprechend perzeptiver Plausibilität der Streichquartett-Konzertsaal-Panoramen die Feinjustierung vorgenommen.

Die Integration aller optischen und akustischen Inhalte mit den fünf stereoskopischen Projektoren und dem an der TU Berlin entwickelten binauralen BK-Kopfhörersystem zu einer hochauflösenden, plausiblen virtuellen Umgebung ermöglicht nun die Untersuchung der Wahrnehmung von akustischen und optischen Räumen von Versuchsteilnehmern. Insbesondere besteht mit dem Medialab an der TU Berlin die Möglichkeit, Konzertsaal-Panoramen und akustische Hallinformationen von Konzertsälen unabhängig voneinander und unabhängig vom Streichquartett zu variieren und damit dem Versuchsteilnehmer verschiedene Kombinationen akustischer, binauraler und optischer, stereoskopischer Räume darzubieten.

Anwendungsbereiche dieser CAVE sind aufgrund der genannten hohen Plausibilität insbesondere in der Perzeptionsforschung, aber auch in der Industrie oder im Marketing vorstellbar. Eine nachhallfreie Audioproduktion und eine stereoskopische Videoproduktion mit Greenscreen-Technik werden in absehbarer Zeit noch dem High End Bereich vorbehalten bleiben.

Mit der Erzeugung des Medialabs am Fachbereich Audiokommunikation an der TU Berlin wurde ein Beitrag für die Weiterentwicklung der Simulationstechnik und der Gestaltung von audio-visuellen, virtuellen Umgebungen geleistet.

Abkürzungen

3D	3 Dimensional
ASIO	A udio S tream I nput/ O utput - standardisiertes Audioübertragungsprotokoll
AVCHD	A dvanced V ideo C odec H igh D efinition - Videoformat mit hoher Auflösung
BRIR	B inaural R oom I mpulse R esponse – binaurale Raumimpulsantwort
CAVE	C ave A utomatic V irtual E nvironment – mittels Computer simulierte künstlich erschaffene (virtuelle) Umgebung
CCD	C harge C oupled D evice – ladungsgekoppeltes Bauelement
CMOS	C ompelementary M etal O xide S emiconductor – Metalloxid-Halbleiter
CODEC	C oder und D ecoder – Verfahren, das Daten digital de/kodiert
CRT	C athode R ay T ube – Kathodenstrahlröhre
DMD	D igital M icromirror D evice – digitales Mikrospiegel Modul
DLP	D igital L ight P rocessing – Projektionstechnologie der Firma Texas Instruments
DSP	D igitaler S ignal P rozessor
FABIAN	F ast and A utomatic B inaural I mpulse response A cquisitio N – schnelle und automatisierte Akquise binauraler Impulsantworten
FEC	F ree air E quivalent C oupling - freifeldäquivalente Ohrkanalankopplung Kopfhörer mit freifeldähnlicher akustischer Abstrahlungsimpedanz
FOV	F ield O f V iew – Bildwinkel
GH	G ewandhaus
HATS	H ead A nd T orso S imulator – Kopf- und Rumpfsimulator
HD	H igh D efinition – hochauflösendes Videoformat
HDTV	H igh D efinition T ele V ision – hochauflösendes Fernsehen
HMD	H ead M ounted D isplay – ein mit dem Kopf verbundenes Display
HpTF	H eadphone T ransfer F unction – Kopfhörer-Übertragungsfunktion
ISO	I nternational O rganization for S tandardization - hier: Filmempfindlichkeitswerte
JCK	J esus- C hristus- K irche
JPEG	J oint P hotographic E xperts G roup – entwickelte diesen Grafik-Standard
KE	K loster E berbach
KH	K onzerthaus
KMQ	K oschnitzke, M ehnert und Q uick – Entwickler eines stereoskopischen Prismenverfahren – des KMQ-Verfahren
KO	K omische O per
LCD	L iquid C rystal D isplay – Flüssigkristallanzeige
LCoS	L iquid C rystal o n S ilicon – Flüssigkristall auf Silikonschicht

LED	L ight E mitting D iode – Licht emittierende Diode
MP	M ega P ixel - Auflösung eines Bildsensors
RGB	R ot G rün B lau – Grundfarben für Erzeugung von Mischfarben
RT	R enaissance- T heater
SD / SDHC	S ecure D igital / S ecure D igital H igh C apacity – Standards für Speicherkarten
SNR	S ignal-to- N oise R atio – Signal-Rausch-Verhältnis
SEACEN	S imulation and E valuation of A coustical E nvironments – Simulation und Evaluation akustischer Umgebungen
TIFF	T agged I mage F ile F ormat – Bilddatenformat
VR	V irtual R eality - virtuelle Realität, virtuelle Umgebung
VRD	V irtual R etinal D isplay – virtuelle Netzhautanzeige, lasergestütztes Abbildungssystem

Quellenverzeichnis

- [1] Eberhardt, Bernhard; Hahn, Jens-Uwe: Computergrafik und Virtual Reality. In: Schmitz, Roland (Hg.): Kompendium Medieninformatik. Berlin Heidelberg, Springer-Verlag, 2007
- [2] Lindau, Alexander: Ein Instrument zur softwaregestützten Messung binauraler Raumimpulsantworten in mehreren Freiheitsgraden. Magisterarbeit, Berlin, TU Berlin, 2006
- [3] Maempel, Hans-Joachim: Audio-visual perception of acoustical environments. Projektantrag, Berlin, TU Berlin, 2011
- [4] Goldstein, E. Bruce: Wahrnehmungspsychologie. Berlin Heidelberg, Springer-Verlag, 2008
- [5] Brill, Manfred: Virtuelle Realität. Berlin Heidelberg, Springer-Verlag, 2009
- [6] Internetquelle: Cruz-Neira, Sandin, DeFanti, Kenyon and Hart, 1992
<http://www.vrarchitect.net/anu/cg/Display/cave.en.html>, Abruf: 26.03.2014
- [7] Internetquelle:http://futurecinema.ca/arlab/wp-content/uploads/2010/07/BNMI100520_DSC9545.jpg, Abruf: 27.03.2014
- [8] Vorländer, Michael u. a.: Virtual Reality System at RWTH Aachen University. In: Proceedings of the International Symposium on Room Acoustics. ISRA 2010 Melbourne, Australia, 29-31 August 2010
- [9] Internetquelle: <http://www.seacen.tu-berlin.de/>. Abruf: 28.03.2014
- [10] Sczepek, Jörg: Visuelle Wahrnehmung. Norderstedt, Books on demand GmbH, 2011
- [11] Internetquelle: Doniec, Marek: Proseminar: Informatik in der Medizin. Universität Karlsruhe, 2003, <http://www.doniec.de/download/S-MDoniec-WS0304.pdf?sid=6487ea5d61d455cc6487ea5d61d455cc2864fffceb82cd31>, Abruf: 29.03.2014
- [12] Anderson, John R.: Kognitive Psychologie. Berlin Heidelberg, Springer-Verlag, 2007
- [13] Baumbach, Ralf: Implementierung einer netzwerkfähigen und interaktiven stereoskopischen Visualisierungsumgebung. Magisterarbeit, Berlin, Technische Universität, 2009
- [14] Hagendorf, Herbert u.a.: Wahrnehmung und Aufmerksamkeit. Allgemeine Psychologie für Bachelor. Berlin Heidelberg, Springer-Verlag, 2011
- [15] Goldstein, E. Bruce: Wahrnehmungspsychologie. Berlin Heidelberg, Springer-Verlag, 2002
- [16] Internetquelle: Hartmann, Axel: Zentralperspektive.
<http://www.flickr.com/photos/axelhartmann/183991009>, Abruf: 03.04.2014
Lizenzbedingungen: <http://creativecommons.org/licenses/by-sa/2.0/deed.de>
- [17] Herbig, Gerhard P.: Die 3 goldenen Regeln der Stereofotografie. 2008,
http://www.cosima-3d.de/download/goldene_regeln_v3.pdf, Abruf: 06.04.2014

- [18] Gockel, Tilo: Kompendium Digitale Fotografie. Berlin Heidelberg, Springer-Verlag, 2012
- [19] Mahler, Gerhard: Die Grundlagen der Fernsehtechnik. Berlin Heidelberg, Springer-Verlag, 2005
- [20] Möser, Michael: Technische Akustik. Berlin Heidelberg, Springer-Verlag, 2009
- [21] Jähne, Bernd: Digitale Bildverarbeitung. Berlin Heidelberg, Springer-Verlag, 2005
- [22] Ansorge, Ulrich; Leder, Helmut: Wahrnehmung und Aufmerksamkeit. Wiesbaden, VS Verlag für Sozialwissenschaften | Springer Fachmedien GmbH, 2011
- [23] Schmidt, Ulrich: Professionelle Videotechnik. Berlin Heidelberg, Springer-Verlag, 2009
- [24] Internetquelle: Wagner, Patrick, Fa. ScanDig: Signalverarbeitung im Gehirn bzw. in der Netzhaut. <http://www.filmscanner.info/Farbwahrnehmung.html>, Abruf: 09.04.2014
- [25] Internetquelle: Funk, Gerhard: Empfindlichkeitsbereich der Stäbchen und Zapfen. <http://www.dma.ufg.ac.at/app/link/Grundlagen%3AAllgemeine/module/16457?step=2>, Abruf: 12.04.2014
- [26] Klein, Georg A: Farbenphysik für industrielle Anwendungen. Berlin Heidelberg, Springer-Verlag, 2004
- [27] Internetquelle: Mischler, Georg: Farbtemperatur. [schorsch.com](http://www.schorsch.com): 2004-2011, <http://www.schorsch.com/de/wissen/glossar/farbtemperatur.html>, Abruf: 12.04.2014
- [28] Böhringer, Joachim; Bühler, Peter; Schlaich, Patrick: Kompendium der Mediengestaltung: Produktion und Technik für Digital- und Printmedien. Berlin Heidelberg, Springer-Verlag, 2011
- [29] Kindermann, Klaus: Fotografieren für Fortgeschrittene: Foto-, Licht- und Aufnahmetechnik optimal einsetzen, Poing, Franzis Verlag, 2008
- [30] Bildquelle: http://www.elearning-psychologie.de/eyetracker_i.html, Abruf: 14.04.2014
- [31] Riedel, Oliver H.: 3D-Echtzeit-Rendering unter Berücksichtigung der Anatomie und Physiologie des menschlichen Auges. Promotionsarbeit, Berlin Heidelberg, Springer-Verlag, 1998
- [32] Böhringer, Joachim; Bühler, Peter; Schlaich, Patrick: Kompendium der Mediengestaltung: Konzeption und Gestaltung für Digital- und Printmedien. Berlin Heidelberg, Springer-Verlag, 2011
- [33] Internetquelle: <http://www.wissenschaft-online.de/abo/lexikon/bio/6051>, Heidelberg, Spektrum Akademischer Verlag, Abruf: 18.04.2014
- [34] Bildquelle: Wähning, Alexander: <http://www.ihr-augenarzt.info/funktion>, Abruf: 20.04.2014
- [35] Internetquelle: <http://www.zeiss.de/C1256B5E00496AB1?Open>, unter: Bilddokumentation ⇒ Digitale Mikrofotografie ⇒ Entstehung eines Bildes, Abruf: 27.04.2014
- [36] Internetquelle: <http://www.zeiss.de/C1256B5E00496AB1?Open>, unter: Bilddokumenta-

- tion ⇒ Digitale Mikrofotografie ⇒ CCD-Sensor ⇒ CCD-Oberfläche, Abruf: 27.04.2014
- [37] Internetquelle: <http://www.zeiss.de/C1256B5E00496AB1?Open>, unter: Bilddokumentation ⇒ Digitale Mikrofotografie ⇒ CCD-Sensor, Abruf: 02.05.2014
- [38] Internetquelle: <http://www.zeiss.de/C1256B5E00496AB1?Open>, unter: Bilddokumentation ⇒ Digitale Mikrofotografie ⇒ Farbdarstellung ⇒ 3-CCD Kamera, Abruf: 05.05.2014
- [39] Internetquelle: <http://www.zeiss.de/C1256B5E00496AB1?Open>, unter: Bilddokumentation ⇒ Digitale Mikrofotografie ⇒ Farbdarstellung ⇒ Bayermaske, Abruf: 05.05.2014
- [40] Internetquelle: <http://www.heise.de/foto/meldung/Kameras-und-Konzepte-Canon-Expo-in-Paris-1107914.html>, Abruf: 07.05.2014
- [41] Schwabe, Martin: Canon EOS 7D. München, Markt+Technik Verlag, 2010
- [42] Internetquelle: <http://fotoseminare.wordpress.com/2011/05/16/film-und-sensorformateil-2-digital/>, Abruf: 08.05.2014
- [43] Schäfer, Claudia: Lichtmikroskopische Untersuchungen zur Rolle des Linsenepithels bei der Kataraktogenese sowie Scheimpflugdokumentation der Linsentrübungsformen bei Typ-II-Diabetikern und Nichtdiabetikern. Dissertation, Martin-Luther-Universität Halle-Wittenberg, 2005, <http://sundoc.bibliothek.uni-halle.de/diss-online/05/05H096/prom.pdf>, Abruf: 14.05.2014
- [44] Internetquelle: http://www.schule-bw.de/unterricht/faecher/physik/online_material/optik/auge/auge_kamera.htm, Landesbildungsserver Baden-Württemberg, Abruf: 16.05.2014
- [45] Dumm, Thomas; Schild, Hansruedi: Strahlenoptik, Schwingungen und Wellen. Zürich, Compendio Bildungsmedien AG, 2009
- [46] Weber, Ernst A.: Fotopraktikum. Birkhäuser Verlag, Basel – Boston – Berlin, 2004
- [47] Internetquelle: <http://www.hcinema.de/kon.htm>, Abruf: 19.05.2014
- [48] Internetquelle: <http://www.hcinema.de/laser.htm>, Abruf: 19.05.2014
- [49] Internetquelle: http://www.chip.de/artikel/LED-Beamer-Konkurrenz-fuer-LCD-und-DLP-Projektor-2_41037507.html, Abruf: 20.05.2014
- [50] Internetquelle: <http://www.dlp.com/de/technology/how-dlp-works>, Abruf: 22.05.2014
- [51] Internetquelle: <http://www.dlp.com/de/technology/dlp-advantages/default.aspx>, Abruf: 22.05.2014
- [52] Internetquelle: <http://electronics.howstuffworks.com/dlp2.htm>, Abruf: 23.05.2014
- [53] Internetquelle: http://www.projector-discount.de/shop_content.php/coID/23/product/Beamer-LCD-vs--DLP, Abruf: 24.05.2014
- [54] Senf, Erhard: Entwicklungsphasen der Stereofotografie. In: Kemner, Gerhard: Stereoskopie. Berlin, Museum für Verkehr und Technik, 1989
- [55] Internetquelle: Ludwig, Michael: http://www.chip.de/artikel/Die-besten-3D-Camcorder-im-Test_50040212.html, Abruf: 27.05.2014

- [56] Internetquelle: <http://www.fotointern.ch/archiv/2010/01/08/panasonic-enthullt-auf-der-ces-2010-den-weltweit-ersten-full-hd-3dcamcorder>, Abruf: 28.05.2014
- [57] Internetquelle: <http://business.panasonic.de/professional-kamera/produkte-und-zubehor/broadcast-und-proav/unser-3d-produktsortiment/ag-3da1>, Abruf: 28.05.2014
- [58] Internetquelle: Rzitki, Fabienne: <http://www.news.de/technik/855216870/shutter-vs-polarisation-3d-technik-im-vergleich/3/>, Abruf: 29.05.2014
- [59] Waack, Fritz G; Kemner, Gerhard: Einführung in Technik und Handhabung der 3-D-Fotografie. In: Kemner, Gerhard: Stereoskopie. Berlin, Museum für Verkehr und Technik, 1989
- [60] Internetquelle: <http://solariz.de/2503/techniken-3d-kino-polarisation-anaglyphisch-shutter.htm>, Abruf: 30.05.2014
- [61] Internetquelle: <http://www.dlp.com/downloads/Introducing%20DLP%203D%20HDTV%20Whitepaper.pdf>, Abruf: 01.06.2014
- [62] Internetquelle: Rzitki, Fabienne: <http://www.news.de/technik/855216870/shutter-vs-polarisation-3d-technik-im-vergleich/2/>, Abruf: 02.06.2014
- [63] Internetquelle: <http://www.infitec.net/index.php/j-stuff/member-access>, Abruf: 03.06.2014
- [64] Internetquelle: Jorke, Helmut; Fritz, Markus: Stereo Projection Using Interference Filters. http://www.infitec.net/images/SPIE_2006.pdf, Abruf: 05.06.2014
- [65] Dultz, Wolfgang; Klein, Susanne: Raumbildtechniken für den Bildschirm. In: Kemner, Gerhard: Stereoskopie. Berlin, Museum für Verkehr und Technik, 1989
- [66] Internetquelle: Schönhals, Johannes: Bilder aus Animation. Digital-Room GmbH, 2011 <http://www.digital-room.de/video/so-funktioniert-3d-ohne-brille>, Abruf: 06.06.2014
- [67] Internetquelle: <https://wiki.engr.illinois.edu/display/BIOE414/TEAM+1+S11+TECHNICAL+DETAILS>, Abruf: 08.06.2014
- [68] Blauert, Jens; Braasch, Jonas: Räumliches Hören. In: Weinzierl, Stefan: Handbuch der Audiotechnik. Berlin Heidelberg, Springer-Verlag, 2008
- [69] Veit, Ivar: Technische Akustik - ausgewählte Kapitel: Grundlagen, aktuelle Probleme und Messtechnik. Renningen, Expert Verlag, 2008
- [70] Schneider, Martin: Mikrofone. In: Weinzierl, Stefan: Handbuch der Audiotechnik. Berlin Heidelberg, Springer-Verlag, 2008
- [71] Internetquelle: <http://www.kimm.uni-luebeck.de/oem/methoden-werkzeuge/mm-audioproduktion/content-6-1.html>, Abruf: 10.06.2014
- [72] Internetquelle: <http://www.kimm.uni-luebeck.de/oem/methoden-werkzeuge/mm-audioproduktion/content-6-2.html>, Abruf: 10.06.2014
- [73] Slavik, Karl M.; Weinzierl, Stefan: Wiedergabeverfahren. In: Weinzierl, Stefan:

- Handbuch der Audiotechnik. Berlin Heidelberg, Springer-Verlag, 2008
- [74] Lindau, Alexander; Weinzierl, Stefan: FABIAN - An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom. Zur 24. TONMEISTERTAGUNG – VDT INTERNATIONAL CONVENTION, November 2006. Berlin, Technische Universität Berlin, Fachgebiet Kommunikationswissenschaft / Audiokommunikation, 2006; Internetquelle: http://www.ak.tu-berlin.de/fileadmin/a0135/Publikationen/2006/Lindau_2006_TMT_FABIAN_An_instrument.pdf, Abruf: 11.06.2014
- [75] Lindau, Alexander; Weinzierl, Stefan: FABIAN – Schnelle Erfassung binauraler Raumimpulsantworten in mehreren Freiheitsgraden. Berlin, Technische Universität Berlin, Fachgebiet Kommunikationswissenschaft / Audiokommunikation, 2007; Internetquelle: http://www2.ak.tu-berlin.de/~akgroup/ak_pub/2007/Lindau_2007b_FABIAN_Schnelle_Erfassung_binauraler_Raumimpulsantworten_DAGA_07.pdf, Abruf: 12.06.2014
- [76] Lindau, Alexander; Maempel, Hans-Joachim: Opto-acoustic simulation of concert halls – a data-based approach (Opto-akustische Simulation von Konzerträumen – ein datenbasierter Ansatz). Zur 27. TONMEISTERTAGUNG – VDT INTERNATIONAL CONVENTION, November 2012. Berlin, Technische Universität Berlin, Fachgebiet Audiokommunikation, 2012; Internetquelle: http://www.ak.tu-berlin.de/fileadmin/a0135/Publikationen/2013/Maempel_2013_Opto-acoustic_simulation_of_concert_halls_-_a_data-based_approach_-_27_TMT_2012.pdf, Abruf: 13.06.2014
- [77] Schule für Rundfunktechnik (Hrsg.), Dickreiter, Michael: Handbuch der Tonstudio-technik, Band1: Raumakustik, Schallquellen, Schallwahrnehmung, Schallwandler, Beschallungstechnik, Aufnahmetechnik, Klanggestaltung. K. G. Saur Verlag, München, 1987
- [78] Erbes, Vera u.a.: An extraaural headphone system for optimized binaural reproduction. In: Fortschritte der Akustik: Tagungsband der 38. DAGA. Fachgebiet Audiokommunikation an der TU Berlin, Berlin, 2012, Internetquelle: http://www.ak.tu-berlin.de/fileadmin/a0135/Publikationen/2012/Erbes_2012_An_extraaural_headphone.pdf, Abruf: 15.06.2014
- [79] Schultz, Frank u.a.: An extraaural headphone for optimized binaural reproduction (Ein extraauraler Kopfhörer für die optimierte Wiedergabe binauraler Signale). Zur 26. TONMEISTERTAGUNG – VDT INTERNATIONAL CONVENTION, November 2010. Berlin, Technische Universität Berlin, Fachgebiet Audiokommunikation, 2010; Internetquelle: http://www2.ak.tu-berlin.de/~akgroup/ak_pub/2011/Schultz_2011_Anextraauralheadphoneforoptimizedbinauralreproduction_26TMT.pdf, Abruf: 17.06.2014

- [80] Internetquelle: http://www.akustik.tu-berlin.de/fileadmin/fg23/Sonstiges/Technische_Einrichtungen/Anechoic_Room.pdf, Abruf: 18.06.2014
- [81] Internetquelle: http://images.thomann.de/pics/atg/atgdata/document/review/184418_test_pa_06_05_kl_einmembran_mikrofone.pdf, Abruf: 19.06.2014
- [82] Weinzierl, Stefan: Aufnahmeverfahren. In: Weinzierl, Stefan: Handbuch der Audiotechnik. Berlin Heidelberg, Springer-Verlag, 2008
- [83] Pawera, Norbert: Mikrofonpraxis. München, PPV Medien, 2004
- [84] Internetquelle: <http://www.genelec.com/documents/datasheets/DS1031a.pdf>, Abruf: 20.06.2014
- [85] Internetquelle: http://www.rme-audio.de/products_fireface_uhx.php, Abruf: 24.06.2014
- [86] Zölzer, Udo: Signalverarbeitung, Filter und Effekte. In: Weinzierl, Stefan: Handbuch der Audiotechnik. Berlin Heidelberg, Springer-Verlag, 2008
- [87] Grimm, Christoph: 3. Vorlesung, Systemtheorie für Informatiker. Frankfurt/Main, Universität. Frankfurt/Main, Internetquelle: http://www.ti.cs.uni-frankfurt.de/lehre/ss04_systemtheorie/v3.pdf, Abruf: 28.06.2014
- [88] Panasonic Corporation: Operating Instruction- Memory Card Camera Recorder AG-3DA1P/1E. Japan 2010, Internetquelle: <http://www.panasonic.com/business/provideo/includes/pdf/AG-3DA1%20OI.pdf>, Abruf: 02.07.2014
- [89] Hurni, Andreas: Ersatzmessung - Die Graukarte. 2014, <http://www.andreashurni.ch/belichtung/messung.htm#ersatzmessung>, Abruf 06.07.2014
- [90] Bourke, Paul: Synthetic Stereoscopic Panoramic Images. In: Zha, Hongbin: Interactive Technologies and Sociotechnical Systems, Berlin Heidelberg, Springer-Verlag, 2006
- [91] Internetquelle: <http://www.ffmpeg.org>, Abruf: 10.07.2014
- [92] Internetquelle: Schmitt, Ekkehart, http://www.cine4home.de/tests/projektoren/Epson_TW9000/TW9000_Test.htm, Abruf: 16.07.2014
- [93] Esswein, Lisa: Aufbau und Konfiguration einer opto-akustischen virtuellen Umgebung. Masterarbeit, Berlin, Technische Universität, 2012

Anhang

A1 Equipmentliste - Audioproduktion Streichquartett

Anzahl	Bezeichnung	Typ
1	Computer	Apple iMac
1	Recordingsoftware	Logic Pro
2	Mikrofon	KM 140
2	Mikrofon	KM 184
1	Mikrofon für Talkback	Sennheiser Evolution E840
2	Aktivmonitor mit Ständer	Genelec 1031A
1	Audiointerface (4xMIC/LINEin)	RME Fireface UFX
1	<i>Audiointerface stereo (2xLINEout, KHout)</i>	?
1	Mischpult (6xMICin, 2xLINEin, 4xAUXpre, 1xAUXpost)	Phonic AM 8440
1	<i>(Hallgerät)</i>	?
2	Kopfhörer	ATH M40-FS
1	Kopfhörer	AKG K271 MKII
1	Kopfhörer	HD 25-1
1	Kopfhörer (Regie)	AKG K 701
1	Kopfhörer (Assistenz)	Sennheiser HD 800
1	Modul mit 4 Kopfhörerverstärkern	Behringer Powerplay Pro HA-4600
1	<i>Einrauscher</i>	?
1	Multicore	10 adrig
5	XLR-Kabel 3m	
8	XLR-Kabel 5m	
6	XLR-Kabel 10m	
4	Verbindung Mixer-Multiface	
1	Verbindung Multiface-Computer	Firewire
4	Kopfhörer-Verlängerungskabel	
2	XLR-Adapter female-female	
2	XLR-Adapter male-male	
4	Klinken-Adapter 6,3female-3,5male	
2	Klinkel-Adapter 3,5female-6,3male	
4	Netzstrom-Verlängerungskabel	
7	Netzstrom-Verteilerdosen (4fach)	
5	Stativ mit Galgen (Mikrofon)	

4	Stativ ohne Galgen (KH-Verstärker)	
8	Stellwand	
7	Stuhl	
5	Notenpulte	
4	Bodenplatten groß (unter Stuhl)	
4	Lampen für Notenpult	
4	Bodenplatte klein (unter Notenpult)	
2	Tischlampe (keine Energiesparbirne!)	
16	Holzleisten	
4	Stoff (akustischer Schirm)	
1	Notenkopien	
1	Takeliste	
1	Equipmentliste	
1	Setup-Plan	
1	Leerpapiere (für Doku)	
2	Klemmbrett	
2	Kugelschreiber	
6	Bleistift	
1	Anspitzer	
6	Radiergummi	
1	Gaffa-Tape	
1	Paketklebeband	
1	Tesafilm	
1	Bindfaden (Rolle)	
7	Wasserflasche 1 Liter (Plastik)	
1	Kl. Schraubenzieher normal	
1	Kl. Schraubenzieher Kreuzschlitz	
1	Lötkolben, Lötzinn, Entlötölitze	
1	Seitenschneider, Justierzange	
1	Cuttermesser	
1	Taschenlampe	
1	Kaffemaschine	
1	Milch	
2	Kekspackung, verschieden	
7	Catering	

A2 Technische Daten der Kleinmembran-Kondensatormikrofone Neumann Km 140 und KM 184

Parameter	KM 140	KM 184
Akustische Arbeitsweise	Druckgrad.- empfänger	Druckgrad.- empfänger
Richtcharakteristik	Niere	Niere
Übertragungsbereich	20 Hz...20 kHz	20 Hz...20 kHz
Feldübertragungsfaktor bei 1 kHz an 1 kOhm	15 mV/Pa	15 mV/Pa
Nennimpedanz	50 Ohm	50 Ohm
Nennlastimpedanz	1 kOhm	1 kOhm
Ersatzgeräuschpegel, CCIR ¹⁾	25 dB	22 dB
Ersatzgeräuschpegel, A-bewertet ¹⁾	16 dB-A	13dB-A
Geräuschpegelabstand, CCIR ¹⁾ (rel. 94 dB SPL)	69 dB	72 dB
Geräuschpegelabstand A- bewertet ¹⁾ (rel. 94 dB SPL)	78 dB	81 dB
Grenzschalldruckpegel für K < 0,5% ²⁾	138 dB	138 dB
Grenzschalldruckpegel für K < 0,5% mit Vordämpfung ²⁾	148 dB	
Max. Ausgangsspannung	10 dBu	10 dBu
Speisespannung (P48, IEC 61938)	48 V ± 4 V	48 V ± 4 V
Stromaufnahme (P48, IEC 61938)	2 mA	3,2 mA
Erforderlicher Steckverbinder	XLR3F	XLR3F
Gewicht	80 g	80 g
Durchmesser	22 mm	22 mm
Länge	92 mm	107 mm

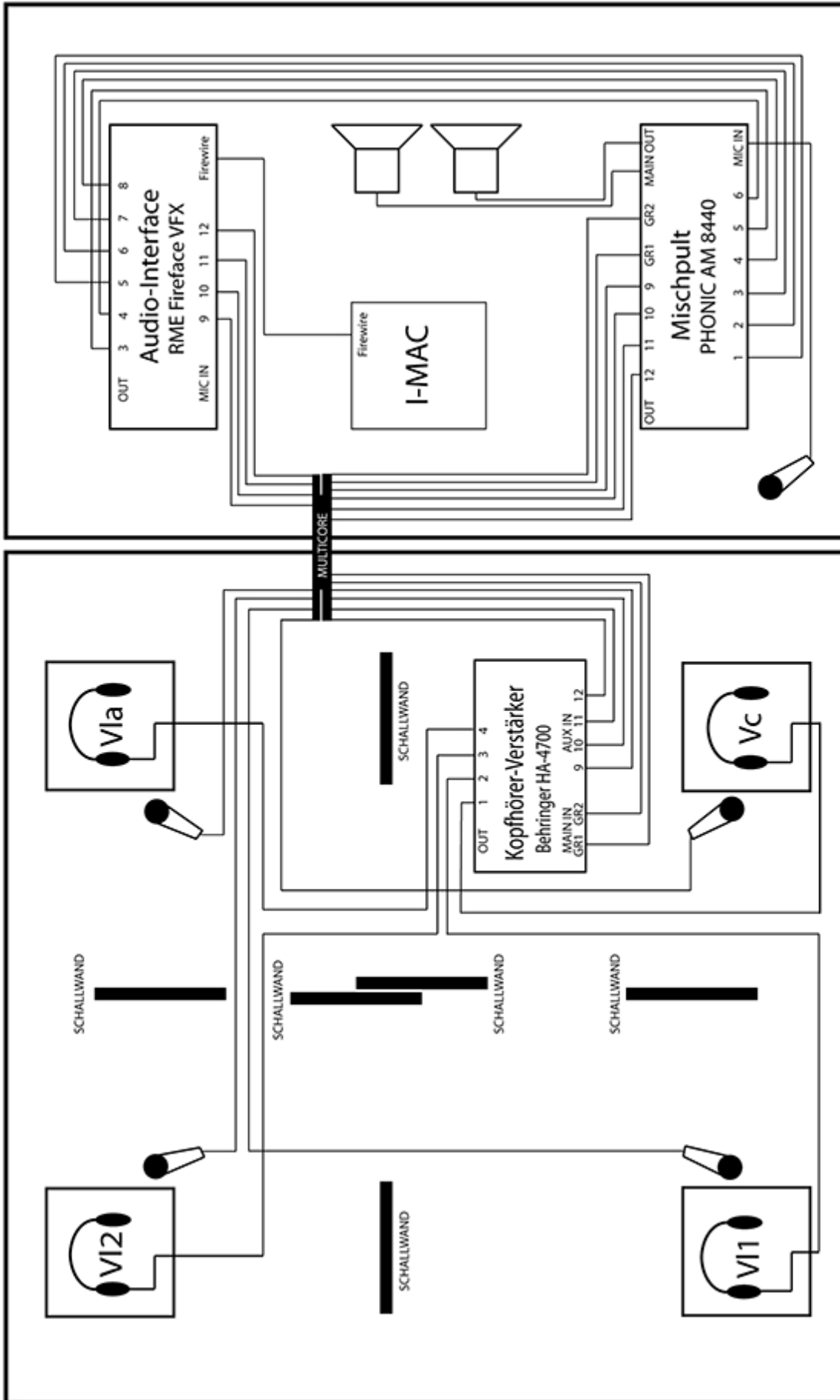
http://www.neumann.com/index.php?lang=de&id=current_microphones&cid=km100_data

http://www.neumann.com/index.php?lang=de&id=current_microphones&cid=km180_data

A3 Mikrofoniertechnik von Violine 2 (VI2) Bratsche (Vla) und Cello (Vc) mit passenden Abstands- und Winkeleinstellungen; vorn und Seite

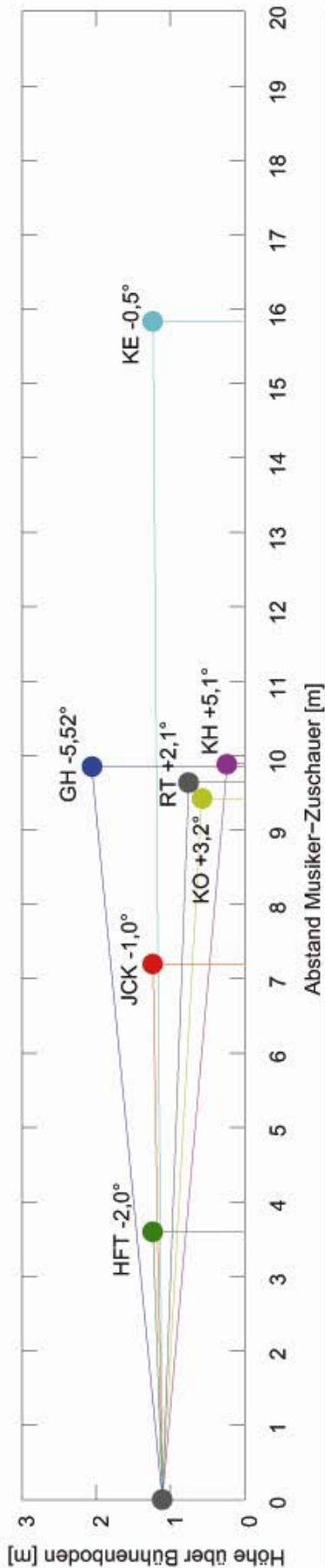


A4 Technischer Aufbau der nachhallfreien Audioproduktion



A5 Ermittlung der Original-Sichtabstände/-Sichthöhen und Aufsichtswinkel

Stereoskopische Greenbox-Videoaufnahme 11.12.2011
 Original-Sichtabstände/-Sichthöhen und Aufsichtswinkel auf ein 1,11m hohes Zentralziel auf der Bühne (Streichquartett-Mitte)



	GH	HFT	JCK	KE	KH	KO	RT
	Gewandhaus Leipzig	Seminarraum HFT-FT617	Jesus-Christus-Kirche Berlin	Kloster Eberbach Basilika	Konzerthaus Berlin Kleiner Saal	Komische Oper Berlin	Renaissance-Theater Berlin
Akquise							
Senderhöhe über Bühnenboden	1,5	1,5	1,5	1,5	1,5	1,5	1,5
Empfängerhöhe (= interaurales Zentrum) über Zuschauerboden*	1,18	1,236	1,236	1,236	1,236	1,236	Steig. 1,48
Bühnenhöhe über Zuschauerboden	-0,8832	0	0	0	0	1	0,9
Senderhöhe über Zuschauerboden	0,6168	1,5	1,5	1,5	1,5	2,5	2,4
Senderhöhe über Zuschauerboden - Empfängerhöhe über Zuschauerboden	-0,5632	0,264	0,264	0,264	0,264	1,264	0,92
Abstand Sender-Empfänger waagrecht [m]	9,86307	3,6	7,2	15,84	15,84	9,888	9,42
Abstand direkt d vom Halsgelenk für 1,5m [m]	9,86	3,61	7,2	15,84	15,84	9,968	9,47
Elevation phi vom Halsgelenk für 1,5m [°]	-3,1	-	-	-	-	6,55	5,8
Elevation phi vom interauralen Zentrum für 1,5m [°]	-3,26815	4,2	2,1	1	1	7,3	5,6
Interaurales Zentrum (= Kamerahöhe) über Bühnenboden	2,0632	1,236	1,236	1,236	1,236	0,236	0,58
Elevation alpha vom interauralen Zentrum für 1,11m [°]	-5,5201	-2	-1	-0,45	-0,45	5,1	3,2
* Sitzhöhe Theaterreihenstuhl 0,42m + Höhe des interauralen Zentrums über Rumpfansatz 0,76m = 1,18m	In-situ-Messung						
	Modellberechnung						

A6 Materialliste für Greenbox-Recording 11.12.2011

Materialliste
2 x Beschallungslautsprecher QSC mit Netzkabel
1 x Laptop ThinkPad mit Netzteil
1 x Audio-Interface Focusrite Saffire 6 mit USB-Kabel
1 x CD-Spieler Marantz professional PMD 321
3 x XLR-Kabel für Zuspeler – Lautsprecher
4 x XLR-Kabel für Lautsprecher – Kamera
2 x Netzkabelverlängerung
1 x Kopfhörer AKG
2 x Netzkabelverlängerung
1 x Kabeltrommel
1 x Netzverteiler
1 x Nikon Kamera von Hanjo
1 x Graukarte / Weißkarte
1 x Target
1 x Lasermeter
1 x Maßband
1 x Mattspray (kauft Mina)
4 x Notenpulte
Audio-Dateien Debussy und Rilke
Audio-CDs Debussy und Rilke
Noten (print)
Excel-Sheet „Kameraperspektiven“ auf Stick
Grafik Sichtabstände (PDF auf Stick / print)
Beispielpanoramen
Wasser
Catering
Andere Getränke

A7 Beleuchtung für Greenbox-Recording 11.12.2011

- 3 x Ianiro "Varibeam" 800Watt
- 2 x Ianiro "Compactbeam" 2000Watt
- 4 x Ianiro "Backlite" 1000Watt (Screenlight)
- 6 x Lowel "Tota-light" 800Watt
- 1 x Domino 1000Watt
- 1 x symmetrischer Fluter 1000Watt
- 1 x Stufenlinse 1000Watt
- 2 x 2000 W Arri mit Chimera & Speedring
- 2 x 800 W Arri Lite inkl. Flügeltore

A8 Brennweite, Belichtungszeiten, Lichtwerte und Motivhelligkeiten bei der Panoramenaufnahme in den Konzertsälen

Motiv	Brennweite [mm]	ISO	Empfindlichkeitsleitwert Sv	Blende f [mm]	Belichtungszeit [s]	Messmethode	Lichtwert (bezogen auf ISO)	Motivhelligkeit Bv, Lv	Motivleuchtdichte L [cd/m ²]
GH_Weißkarte	18	200	6	3,5	0,10000	Mittenbetont (!)	6,94	0,94	7,13
GH_Graukarte	18	200	6	3,5	0,40000	Mittenbetont (!)	4,94	-1,06	1,78
GH_Bühnenboden	18	200	6	3,5	0,04000	Mittenbetont (!)	8,26	2,26	17,80
GH_Bühnenrückwand Holz	18	200	6	3,5	0,40000	Mittenbetont (!)	4,94	-1,06	1,78
GH_Bühnenrückwand Gips	18	200	6	3,5	0,04000	Mittenbetont (!)	8,26	2,26	17,80
HFT_Weißkarte	18	200	6	3,5	0,02500	Spot	8,94	2,94	28,52
HFT_Graukarte	18	200	6	3,5	0,10000	Spot	6,94	0,94	7,13
HFT_Tafelecke links oben (dunkel)	18	200	6	3,5	0,25000	Spot	5,62	-0,38	2,86
HFT_Fußboden vor Tafel (dunkel)	18	200	6	3,5	0,01667	Spot	9,52	3,52	42,63
HFT_Graue Tür (mittel)	18	200	6	3,5	0,07692	Spot	7,32	1,32	9,28
JCK_Weißkarte	18	200	6	3,5	0,05000	Spot	7,94	1,94	14,26
JCK_Graukarte	18	200	6	3,5	0,25000	Spot	5,61	-0,39	2,84
JCK>Weisse Seitenwand links über Spieltisch (mittel)	18	200	6	3,5	0,05000	Spot	7,94	1,94	14,26
JCK_Fußboden links vor Querläufer (hell)	18	200	6	3,5	0,01250	Spot	9,94	3,94	57,04
JCK_Altartuch grün (dunkel)	18	200	6	3,5	0,50000	Spot	4,61	-1,39	1,42
JCK_Tür links (dunkel)	18	200	6	3,5	0,33333	Spot	5,20	-0,80	2,13
KE_Weißkarte	18	200	6	3,5	0,04167	Spot	8,20	2,20	17,08
KE_Graukarte	18	200	6	3,5	0,20000	Spot	5,94	-0,06	3,56
KE_Altarfenster (dunkel)	18	200	6	3,5	2,66667	Spot	2,20	-3,80	0,27
KE_Bühnenboden (mittel)	18	200	6	3,5	0,61538	Spot	4,32	-1,68	1,16
KE_Seitenschiffwand	18	200	6	3,5	0,04667	Spot	8,04	2,04	15,28

(hell)									
KH_Weißkarte	18	200	6	3,5	0,00800	Spot	10,58	4,58	88,89
KH_Graukarte	18	200	6	3,5	0,05000	Spot	7,94	1,94	14,26
KH_Bühnenboden (hell)	18	200	6	3,5	0,01000	Spot	10,26	4,26	71,20
KH_Bühnenrückwand mitte (mittel)	18	200	6	3,5	0,10000	Spot	6,94	0,94	7,13
KH_Bühnenrückwand unten (dunkel)	18	200	6	3,5	0,16667	Spot	6,20	0,20	4,27
KO_Weißkarte auf Bühne	18	200	6	3,5	0,00250	Spot	12,26	6,26	284,82
KO_Graukarte auf Bühne	18	200	6	3,5	0,01000	Spot	10,26	4,26	71,20
KO_Weißkarte im Parkett	18	200	6	3,5	0,03333	Spot	8,52	2,52	21,32
KO_Graukarte im Parkett	18	200	6	3,5	0,16667	Spot	6,20	0,20	4,27
KO_Linkes mittleres Proszeniumsfenster (hell)	18	200	6	3,5	0,05000	Spot	7,94	1,94	14,26
KO_Mitte des roten Vorhangs (mittel)	18	200	6	3,5	0,25000	Spot	5,61	-0,39	2,84
KO_Linke Proszeniumsloge (dunkel)	18	200	6	3,5	5,00000	Spot	1,29	-4,71	0,14
RT_Weißkarte	18	200	6	3,5	0,02500	Spot	8,94	2,94	28,52
RT_Graukarte	18	200	6	3,5	0,07692	Spot	7,32	1,32	9,28
RT_Bühnenboden (hell)	18	200	6	3,5	0,01667	Spot	9,52	3,52	42,63
RT_Bühnenrückwand Mitte (dunkel)	18	200	6	3,5	4,00000	Spot	1,61	-4,39	0,18
RT_Holzvertäfelung Seite rechts (mittel)	18	200	6	3,5	1,60000	Spot	2,94	-3,06	0,45
									Bv=0 entspricht 4 cd/qm
									Bv=ld(L)-2
									$L=2^{(Bv+2)}$

A9 Trapezkorrektur - Skalierungen

Saal	Perspektive	Trapezkorrektur (Zielwerte)		Gesamtskalierung (Faktor)
		Obere Bildbreite	Untere Bildbreite	
GH	C	1922 Pixel	1918 Pixel	59,84 %
JCK	B	1922 Pixel	1918 Pixel	67,95 %
KE	B	1922 Pixel	1918 Pixel	30,89 %
KH	A	1934 Pixel	1910 Pixel	37,79 %
KO	A	1921 Pixel	1919 Pixel	39,67 %
RT	A	1920 Pixel	1920 Pixel	38,76 %

A10 Matlab-Script „Bilder_zusammensetzen.m“ für das Zusammenfügen der Streichquartett-Bilder mit den Panoramen

Matlab-Script „Bilder_zusammensetzen.m“:

```

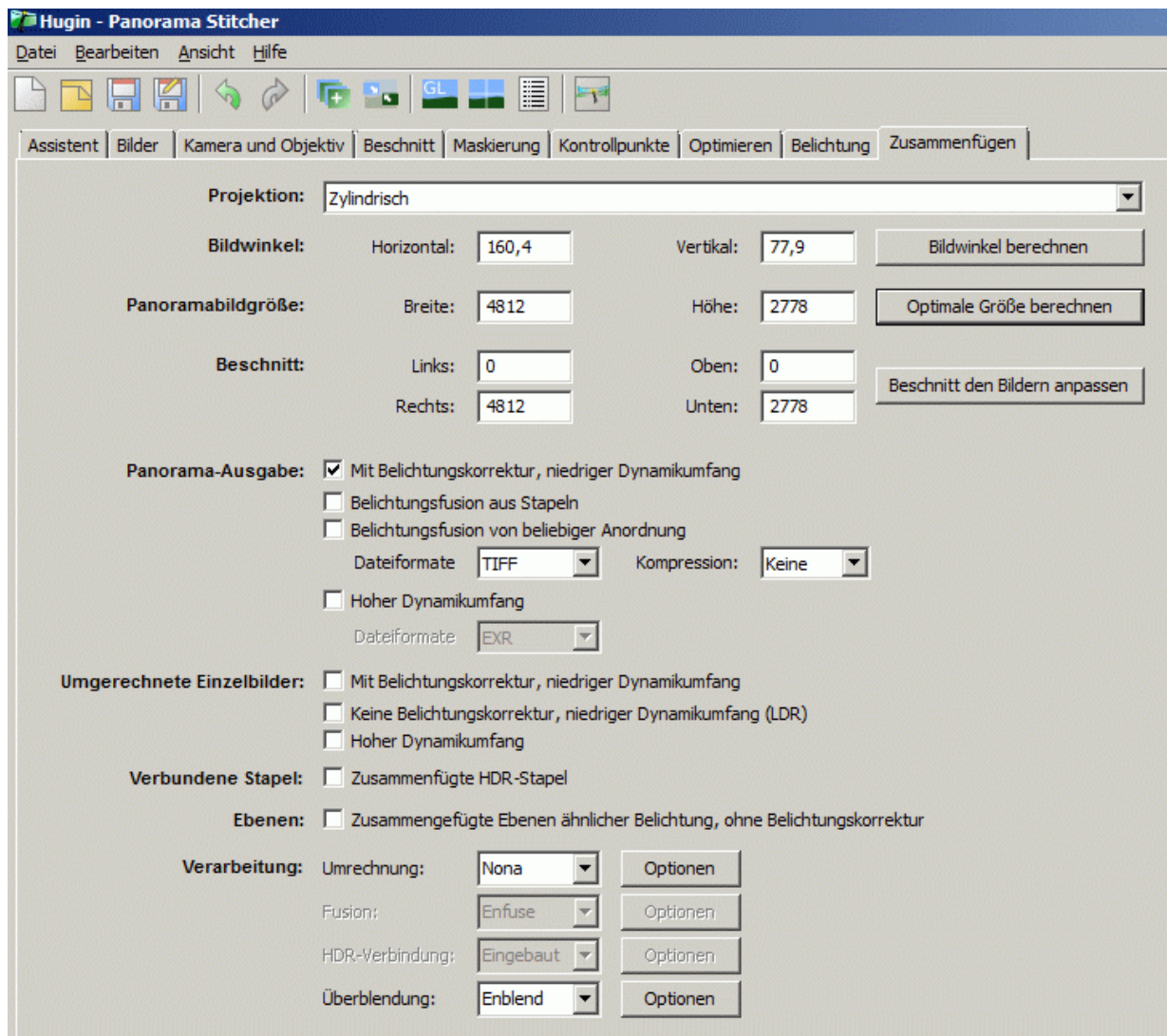
close all;
clear all;
clf;
clc;

% Folgende Schleifen sind für 10 Sekunden. Jede Sekunde beinhaltet 30
% Frames:

for j = 1:10
    for i = 0:29
        cd('E:\Mina\GH_Link_rücktransfromiert');
        filenameleft=['GH_L_' num2str(j, '%03d') '_' num2str(i,'%04d') '.tif'];
        A =
imread('GH_panoEqui_cutout180°x90°symmHorizont_L_centered_verticallyAligned_resized
5400x2700.tif');
        B = imread(filenameleft);
        C2=A;
        C2(2300:3099,2000:3399,:)= B(140:939,260:1659,1:3); % 800*1400 Pixel von Bild C
werden
        % mit Pixeln vom Bild B ersetzt.
        cd('E:\Mina\GH_R_Pano_zuschnitt');
        filename2=['GH_L_' num2str(j, '%03d') '_' num2str(i,'%04d') '_pano.tif']
        % Das Ergebnis speichern:
        imwrite(C2, filename2);
        close all;
        end
    end
end

```

A11 Konvertierung der Panorama-Frames von einer equirektangularen in eine zylindrische Projektionsgeometrie



Abbildungsverzeichnis

Abbildung 1: Prinzip der CAVE an der TU Berlin	2
Abbildung 2: Untersuchung cross modaler Effekte	2
Abbildung 3: Wahrnehmungsprozess	4
Abbildung 4: drei Teilschritte des Wahrnehmungsprozesses	4
Abbildung 5: Prinzip-Aufbau einer CAVE	5
Abbildung 6: in einer CAVE	5
Abbildung 7: Grundriss der CAVE am Fachbereich Audiokommunikation der TU Berlin	6
Abbildung 8: Anatomie des menschlichen Auges	7
Abbildung 9: Objektabbildung auf der Netzhaut; Rezeptoren in der Netzhaut	8
Abbildung 10: Schnitt durch das Gehirn mit den wichtigsten Stationen der Sehbahnen	9
Abbildung 11: Aufbau der Sehrinde	9
Abbildung 12: Überlagerung von linker und rechter Netzhaut	10
Abbildung 13: gekreuzte und ungekreuzte Querdisparation	10
Abbildung 14: Tiefenkriterien der Raumwahrnehmung	12
Abbildung 15: Konvergenzwinkel in Abhängigkeit von der Objektentfernung	12
Abbildung 16: Anpassung des Auges an unterschiedliche Objektentfernung	12
Abbildung 17: Verdeckung, relative Größe, relative Höhe, Perspektive	14
Abbildung 18: atmosphärische Perspektive	14
Abbildung 19: Zentralperspektive	14
Abbildung 20: Bewegungsparallaxe	15
Abbildung 21: fortschreitendes Zu- und Aufdecken von Flächen	15
Abbildung 22: Wellenlängen des sichtbaren Lichtes	17
Abbildung 23: spektrale Hellempfindlichkeitskurven für das Stäbchen- und Zapfensehen	18
Abbildung 24: Verschaltung von Stäbchen und Zapfen	18
Abbildung 25: Qualitativer Zusammenhang zwischen Reiz und Empfindung	20
Abbildung 26: simultaner Helligkeitskontrast	20
Abbildung 27: relative Empfindlichkeit der drei Zapfentypen	21
Abbildung 28: Erfassung und Vorverarbeitung des Lichtreizes	21
Abbildung 29: Einordnung natürlicher und künstlicher Lichtquellen	22
Abbildung 30: Eyetracker	23
Abbildung 31: individueller Blickbewegungsverlauf	23
Abbildung 32: Gesichtsfeld des Menschen	24
Abbildung 33: horizontales und vertikales Gesichtsfeld des Menschen	24
Abbildung 34: Örtliches Auflösungsvermögen des Auges	25
Abbildung 35: Strahlengang und Bildentstehung bei Auge und Kamera	27

.Abbildung 36: Bayer-Matrix.....	29
.Abbildung 37: 3CCD Technik.....	29
.Abbildung 38: Funktionsprinzip eines CMOS- Sensors.....	30
.Abbildung 39: Strahlengang und Bildentstehung durch Sammellinse.....	31
Abbildung 40: Bildwinkel und Brennweiten unterschiedlicher Objektivarten.....	33
Abbildung 41: Schärfentiefebereich in Abhängigkeit von der Blende.....	33
.Abbildung 42: Vergleich von Auge und Kamerasystemen.....	35
.Abbildung 43: Übersicht über Bildwiedergabewandler.....	36
.Abbildung 44: Prinzip eines CRT-Projektors.....	37
.Abbildung 45: Prinzip eines Laser-Projektors.....	37
.Abbildung 46: Prinzip eines DLP-Projektors.....	39
.Abbildung 47: Prinzip eines LCD-Projektors.....	39
.Abbildung 48: stereoskopische Aufnahme- und Wiedergabeverfahren.....	40
.Abbildung 49: optische Anordnungen von Side by Side- und Mirror-Rigs.....	42
.Abbildung 50: Panasonic AG-3DA13D; 3D-Camcorder mit integriertem Doppelobjektiv.....	42
.Abbildung 51: Anaglyphenbild.....	43
.Abbildung 52: Polarisationsverfahren, Prinzip.....	43
Abbildung 53: Shutterverfahren, Prinzip.....	44
Abbildung 54: Wellenlängenmultiplex Verfahren, Prinzip.....	46
Abbildung 55: Virtual Retinal Displays, Prinzip.....	46
Abbildung 56: Autostereoskopie; Parallaxenbarriere.....	47
Abbildung 57: Autostereoskopie; Linsenraster für mehrere Betrachter.....	47
Abbildung 58: Zeit- und Frequenzverhalten eines einfachen Klages.....	49
Abbildung 59: Hörfläche, Hörschwellenkurve und Föhlschwelle.....	49
Abbildung 60: Das Ohr des Menschen.....	50
Abbildung 61: Cochlea.....	50
Abbildung 62: Querschnitt durch die Cochlea; Corti'sches Organ.....	51
Abbildung 63: Tonotope Karte der menschlichen Cochlea.....	52
Abbildung 64: Cochlea zerlegt Schallereignisse in Frequenzanteile.....	52
Abbildung 65: interaurale Zeitdifferenz.....	53
Abbildung 66: interaurale Pegeldifferenz.....	53
Abbildung 67: Primärschall im Freifeld.....	55
Abbildung 68: Primärschall und reflektierter Schall in geschlossenen Räumen.....	55
Abbildung 69: Lokalisation von Schallereignissen.....	56
Abbildung 70: Klassifizierung von Luftschallmikrofonen.....	58
Abbildung 71: dynamisches Mikrofon, Prinzip.....	59
Abbildung 72: Kondensatormikrofon, Prinzip.....	59

Abbildung 73: Binauraltechnik: Aufnahme der BRIRs mit Kunstkopf	60
Abbildung 74: Messsystem FABIAN zur Akquise binauraler Raumimpulsantworten	60
Abbildung 75: binauraler Kopfhörer Prototyp BK109	63
Abbildung 76: Frequenzgang des binauralen Kopfhörers BK109	63
Abbildung 77: Konzept der Reizerzeugung für die SEACEN-CAVE	64
Abbildung 78: reflexionsarmer Raum des Instituts für Technische Akustik	65
Abbildung 79: Stellwände im reflexionsarmen Raum	65
Abbildung 80: Frequenzgänge des Aufnahmемikrofons KM 184	67
Abbildung 81: Mikrofon-Instrumenten-Abstände, Einfalls- und Abstrahlwinkel	68
Abbildung 82: Mikrofonierung von Violine 1; vorn	69
Abbildung 83: Mikrofonierung von Violine 1; Seite	69
Abbildung 84: Kopfhörerverstärker Behringer Power-play Pro HA-4600	70
Abbildung 85: Regieplatz zum Abhören der Aufnahmen im reflexionsarmen Raum	70
Abbildung 86: Technischer Aufbau der nachhallfreien Audioproduktion	71
Abbildung 87: Positionen der einzelnen Streichinstrumente bei der Akquise der BRIRs ...	72
Abbildung 88: 3D Bild des in Greenscreenbox spielenden Streichquartetts	74
Abbildung 89: Effekt der Trapezverzerrung	75
Abbildung 90: Streichquartett-Aufsichtwinkel gegenüber dem Konzertsaalhintergrund	75
Abbildung 91: Bestimmung der Streichquartett-Aufsichtwinkel	76
Abbildung 92: Tabelle zur Bestimmung der Aufsichtwinkel	76
Abbildung 93: grafische Erfassung der Aufsichtwinkel aller Konzertsäle	77
Abbildung 94: Tabelle zur Bestimmung der Aufnahmehöhe	77
Abbildung 95: Greenscreen-Studio für Streichquartett	79
Abbildung 96: 3D Camcorder Panasonic AG-3DA1	79
Abbildung 97: Frontansicht des verwendeten Panasonic AG-3DA1 Camcorders	82
Abbildung 98: Bestimmung des Parallaxenwinkels	82
Abbildung 99: Beleuchtung bestehend aus Strahlern, Flutern etc.	83
Abbildung 100: Berlin-Budapest Quartett spielt vor Greenscreen-Wand	83
Abbildung 101: Aufnahme der drei Perspektiven	84
Abbildung 102: Bilder der sechs Hintergrundpanoramen	85
Abbildung 103: digitale Bildverarbeitungsstrecke	85
Abbildung 104: Gewandhaus Leipzig; equirektangularer HD-Panoramaausschnitt	86
Abbildung 105: Gewandhaus Leipzig; rectilinearer HD-Panoramaausschnitt	86
Abbildung 106: Bildbetrachtung mit Aufsichtwinkel	87
Abbildung 107: Montage von rectilinearen Streichquartett-Aufnahmen	89
Abbildung 108: Konvertierung der Streichquartett-Panoramaausschnitt-Frames	89
Abbildung 109: Halbpanorama des Gewandhauses; equirektangular	90

Abbildung 110: Halbpanorama des Gewandhauses mit Panoramaausschnitt-Frame	90
Abbildung 111: CAVE – Seitenansicht.....	91
Abbildung 112: CAVE - Draufsicht.....	91
Abbildung 113: Streichquartett-Konzertsaal-Panorama, zylindrisch.....	93
Abbildung 114: Leistungsverstärker vom Typ <i>BKamp</i>	96
Abbildung 115: binauraler Kopfhörer vom Typ <i>BK 211</i>	96
Abbildung 116: CAVE im Medialab an der TU Berlin.....	97
Anhang A2: Technische Daten der Mikrofone Neumann Km 140 und KM 184.....	110
Anhang A3: Mikrofonierung von Violine 2 (Vl2) Bratsche (Vla) und Cello (Vc).....	111
Anhang A4: Technischer Aufbau der nachhallfreien Audioproduktion.....	112
Anhang A5: Ermittlung der Original-Sichtabstände/-Sichthöhen und Aufsichtswinkel.....	113
Anhang A11: Konvertierung der Panorama-Frames.....	120