

Eigenschaften und Kriterien für gute binaurale Aufnahmesysteme

Bachelorarbeit aus Aufnahmetechnik 1, SE

Konstantin Kaltenegger

Betreuung: Univ.Prof. DI Dr. Alois Sontacchi

Graz, 31. August 2022



institut für elektronische musik und akustik



Zusammenfassung

Die vorliegende Arbeit gibt einen Überblick in das Themengebiet der binauralen Aufnahmetechnik. Im Speziellen wird darauf eingegangen welche Qualitätskriterien ein qualitativ hochwertiges Aufnahme-Wiedergabesystem erfüllen sollte. Dabei wird kurz das theoretische Grundwissen im Gebiet des räumlichen Hörens erläutert um in weiterer Folge die Eigenschaften binauraler Aufnahmesysteme besser bewerten zu können. Anschließend folgt ein Überblick über die Technologie der kopfbezogenen Stereophonie. Es werden Vor- bzw. Nachteile der Kunstkopftechnik sowie der Echkopftechnik vorgestellt und anschließend diskutiert.

Inhaltsverzeichnis

1	Einleitung	5
2	Historischer Überblick über die Entwicklung von Kunstköpfen	6
3	Räumliches Hören	8
3.1	Kopfbezogenes Koordinatensystem	8
3.2	Richtungslokalisation in der Horizontalebene	9
3.2.1	Interaurale Laufzeitdifferenz	9
3.2.2	Interaurale Pegeldifferenz	9
3.2.3	Duplex Theorie	9
3.2.4	Trading Effekt	10
3.3	Richtungslokalisation in der Medianebene	11
3.3.1	Eigenschaften des Außenohres	11
3.3.2	Blauert'sche Bänder	12
3.4	Außenohrübertragungsfunktion	12
3.5	Entfernungswahrnehmung	13
3.6	Lokalisation von mehreren Schallquellen	14
4	Binaurale Aufnahmesysteme	16
4.1	Prinzip von binauralem Stereo	16
4.2	Aufbau des Kunstkopfes	16
4.3	Echtkopfmikrofonierung als Alternative zum Kunstkopf	18
4.4	Prinzip der Aufnahme- und Wiedergabekette	19
4.5	Anwendungsgebiete	21
4.6	Herausforderungen bei binauralen Systemen	22
4.7	Anforderungen an ein hochwertiges Aufnahmesystem	24
4.8	Vergleich von Messungen realer Systeme	25

<i>K. Kaltenegger: Eigenschaften guter binauraler Aufnahmesysteme</i>	4
5 Diskussion und Ausblick	29

1 Einleitung

Durch die Fähigkeit Schallereignisse in natürlicher Hörumgebung dreidimensional wahrzunehmen ist der Mensch in der Lage sich in seinem Umfeld zu orientieren. Dieses akustische Umfeld wird von unseren beiden Ohren erfasst, übertragen, in neuronale Reize umgewandelt und von unserem Gehirn verarbeitet. Durch diese Übertragungskette sind wir in der Lage eintreffenden Schallereignissen eine Richtung zuzuweisen und auf diese zu reagieren. Wie dieses natürliche Hörerlebnis nun eingefangen und reproduziert werden kann, beschäftigt Forscher seit vielen Jahren. Ein Weg dies zu bewerkstelligen ist unter anderem der Einsatz von binauralen Aufnahmesystemen. Im Idealfall sollte durch die bestmögliche Aufnahme sowie Wiedergabe der hörenden Person ein Hörerlebnis wie in natürlicher Hörumgebung geboten werden. Im Verlauf dieser Arbeit wird auf die Kriterien und Problemstellungen eingegangen, welche beim Entwurf eines guten binauralen Aufnahmesystems zu beachten sind.

Um besser nachvollziehen zu können welche Evolutionsschritte nötig waren um die Technologie der binauralen Aufnahmesysteme weiterzuentwickeln wird in Kapitel 2 ein kurzer historischer Überblick über die Meilensteine in der Entwicklung solcher Systeme gegeben. In Kapitel 3 wird näher auf die akustischen, anatomischen und psychoakustischen Mechanismen eingegangen, die hinter dem Phänomen des räumlichen Hörens stehen. Hierbei wird im speziellen auf die Richtungs- sowie die Entfernungslokalisation eingegangen, da das Verständnis in diesen Gebieten in weiterer Folge wichtig für die Beschreibung der Eigenschaften binauraler Aufnahmesysteme ist.

Das Kapitel 4 beschäftigt sich mit den binauralen Aufnahmesystemen, sowie mit damit verbundenen Aspekten, die bei der Analyse sowie der Konstruktion von Systemen beachtet werden sollten. Dabei wird Hauptaugenmerk auf den Kunstkopf gelegt, allerdings werden auch Alternativen wie die Echkopf-Aufnahme besprochen. Es werden Vor- bzw. Nachteile der beiden Ansätze genannt und in Folge dessen werden diese Aussagen mit Ergebnissen wissenschaftlicher Studien verglichen und in Kapitel 5 diskutiert.

2 Historischer Überblick über die Entwicklung von Kunstköpfen

Die binaurale Aufnahmetechnik ist kein Phänomen, das sich in wenigen Jahren entwickelt hat. Forscher, Erfinder und Ingenieure sind seit über 100 Jahren bemüht einen Weg zu finden das menschliche Hörerlebnis originalgetreu aufzunehmen, zu übertragen und zu reproduzieren. Dieses Kapitel gibt einen kurzen Überblick über die Meilensteine die zum heutigen Stand der Kunstkopftechnologie geführt haben.

Der Begriff *binaural* wird bereits 1861 erstmals in der Literatur von Alison verwendet um das Hören mit zwei Ohren zu beschreiben. Damals verwendete man das Wort allerdings noch nicht im Bezug auf die, durch den menschlichen Kopf, Körper sowie durch das Außenohr modifizierten Ohrsignale. Außerdem gab es noch keine strenge Trennung zum Begriff *Stereo*. [Pau09]

Der initiale Schritt zur paarweisen Stereomikrofontechnik wurde 1881 von Clement Ader gesetzt, welcher eine Opernaufführung auf zwei Kanälen in entfernte Räumlichkeiten übertragen ließ. Hierfür installierte Ader jeweils paarweise zwei Mikrofone an den linken bzw. rechten Bühnenrand und übertrug die Signale an zwei Ohrmuscheln, die in zwei Kilometer Entfernung von den Zuhörer*Innen getragen wurden. Diese Erfindung trägt den Namen „Theatrophon“ und kann als das erste binaurale Aufnahmesystem gesehen werden.

Zum ersten nennenswerten Einsatz einer Puppe mit eingebauten Mikrofonen kam es um 1930 in den Bell Telephone Laboratories in Murray Hill, in den Vereinigten Staaten. Mit der auf Abbildung 1 dargestellten Wachspuppe *Oscar* wurde zu dieser Zeit experimentiert um räumliches Hören besser zu verstehen. [WBG⁺17]

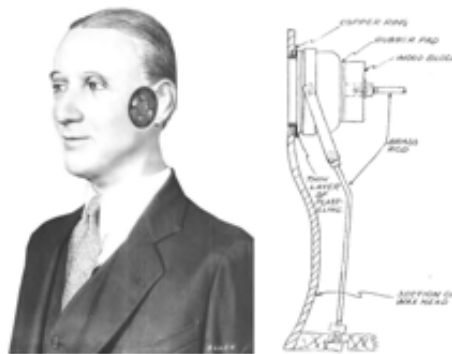


Abbildung 1 – "Oscar" entwickelt von Fletcher an den Bell Laboratories [Pau09]

Bis zum nächsten Evolutionsschritt der Kunstkopftechnologie dauerte es dann bis in die späten 1960er Jahre. Eine Forschungsgruppe um Georg Plenge, Ralf Kürer, und Henning Wilkens baute in Deutschland ein weiteres Kunstkopfsysteme, um die Technologie für wissenschaftliche Studien im Bereich der Raumakustik zu verwenden. Hierfür wurde zum ersten Mal das menschliche Außenohr inklusive dem Gehörgang, nachgebildet - siehe Abbildung 2. [TMH20]

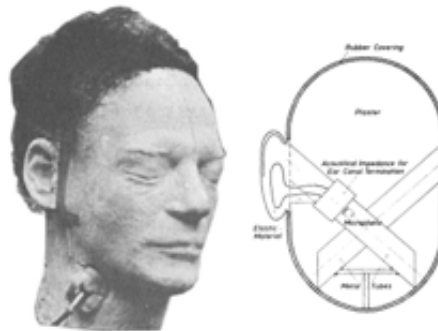


Abbildung 2 – Der erste im Jahr 1969 entwickelte "Kunstkopf" [Pau09]

Nachdem die Wissenschaftler zwei Mikrofone des Typs KU83 der Firma Neumann im Inneren des Kunstkopfes verbaut hatten, entstand eine Kooperation mit der Firma die im Jahr 1972 zum ersten kommerziell produzierten Kunstkopfsystem - dem *Neumann KU80* (siehe Abbildung 3) - führte. [TMH20]

Das erste binaurale Hörspiel "Demolition" welches 1973 an der Internationalen Radio Exhibition in Berlin vorgestellt wurde war ein großer Erfolg, trotzdem gelang es der Technologie nicht zu kommerziellem Erfolg. In den weiteren Jahren wurde die Kunstkopftechnologie noch stetig weiter verbessert, doch das grundlegende Fundament der Technologie war bis zu diesem Zeitpunkt bereits gelegt.



Abbildung 3 – 1973 brachte die Firma Neumann den ersten kommerziellen Kunstkopf KU80 heraus [Pau09]

Es folgten weitere Modelle der Firmen Sennheiser, AKG, Neumann, Kemar uvm. welche die Technologie weiter, mit Hoffnung auf kommerziellen Erfolg, vorantrieben. In dieser Hinsicht setzte sich die binaurale Aufnahmetechnik bis heute nicht durch, allerdings ist sie, wie in weiterer Folge in dieser Arbeit berichtet wird, ein zuverlässiges Mittel um einen besseren Einblick in die auditive Wahrnehmung, also in die Sinneswahrnehmung des Menschen, zu bekommen.

3 Räumliches Hören

Um den Begriff des "Räumlichen Hörens" zu beschreiben, gibt es unterschiedliche Ansätze, welche sowohl physiologischer als auch psychologischer Natur sein können. Ganz allgemein lässt sich jedoch behaupten, dass man darunter die Beziehungen zwischen dem Ort sowie der räumlichen Ausdehnungen eines Hörereignisses und den damit einhergehenden Merkmalen versteht. [BB08]

Im folgenden Kapitel wird ein kurzer Überblick über die wichtigsten Merkmale gegeben, welche die Grundlage des räumlichen Hörens darstellen. Dabei wird besonderes Augenmerk auf die Lokalisationsfähigkeit des Menschen gelegt, da diese einen wichtigen Aspekt für die Analyse von binauralen Aufnahmesystemen darstellt.

3.1 Kopfbezogenes Koordinatensystem

Um den Hörereignissort bzw. die räumliche Ausdehnung einer Schallquelle beschreiben zu können wird das kopfbezogene Koordinatensystem verwendet. Dieses leitet sich vom *räumlichen Polarkoordinatensystem* ab, welches zur Beschreibung eines Ursprungspunktes im dreidimensionalen Raum mit Hilfe zweier Winkel sowie dem Abstand zum Ursprung beschrieben wird. Den Ursprung stellt hier der geometrische Mittelpunkt des Kopfes dar, welcher auf der interauralen Achse in gleichem Abstand zu den beiden Ohrkanaleingängen liegt. [BB08]

Wie in Abbildung 4 zu sehen ist, lässt sich der Raum in drei Ebenen unterteilen - die *Medialebene* (Sagittalebene), die *Frontalebene* sowie die *Horizontalebene* (Transversalebene). Bei der Koordinierung auf diesen Ebenen wird der Abstand r von der hörenden Person zur Schallquelle, der Azimut-Winkel φ , sowie der Elevations-Winkel δ genutzt.

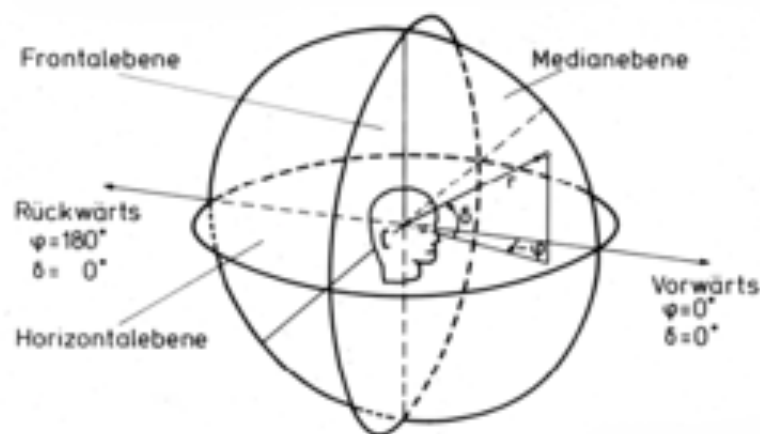


Abbildung 4 – Kopfbezogenes Koordinatensystem [Bra02]

3.2 Richtungslokalisierung in der Horizontalebene

Bei der Richtungslokalisierung des Menschen muss je nach Einfallswinkel zwischen verschiedene Wahrnehmungsmodellen unterschieden werden. Auf der Horizontalebene wird binaurales Hören benötigt, um *interaurale Laufzeitdifferenzen* (engl. interaural time difference, ITD), sowie *interaurale Pegeldifferenzen* (engl. interaural level difference, ILD) wahrnehmen zu können. Hierfür sind, wie es das Wort schon vermuten lässt, zwei Empfänger (Ohren) nötig, mit denen es möglich ist Informationen über die Richtung des Schalleignisses zu gewinnen. Auf diese beiden Konzepte, sowie mit ihnen verbundene Effekte wird in den folgenden Unterpunkten näher eingegangen.

3.2.1 Interaurale Laufzeitdifferenz

Eine Klangquelle, welche bezüglich des azimuthalen Winkels von 0° bzw. 180° abweicht wird gezwungenermaßen eine gewisse interaurale Laufzeitdifferenz aufweisen. Durch die Wegdifferenz, die sich durch die Ohrpositionen ergibt kommt das Quellsignal an einem der beiden Ohren zeitverzögert an. Da der Schall sich in der Luft mit etwa 340 m/s ausbreitet und die durchschnittliche Kopfbreite eines Menschen häufig mit etwa 22 cm angenommen wird, ergibt sich dadurch etwa eine maximale ITD von 0,79 ms. [Rum01]

Diese ITD tritt bei einem Schalleinfallswinkel von 90° Azimut auf, da der Schall einen zusätzlichen Weg zurücklegen muss um das weiter entfernte Ohr zu erreichen. In Studien wurde mehrfach gezeigt, dass der Mensch mit Hilfe dieser Methode in der Lage ist einen MAA (engl. Minimum Audible Angle) von bis zu 1° wahrzunehmen, wenn die Schallquellen lediglich auf der Horizontalebene platziert werden (dh. wenn nur der Azimut für die Quellenortung genutzt wird). In diesen Tests kann die ITD allerdings niemals isoliert betrachtet werden, da die ILD hier die Schallquellenortung beeinflusst. [PS90]

3.2.2 Interaurale Pegeldifferenz

Schallwellen, welche eine kleinere Wellenlänge, als die Abmessung des menschlichen Kopfes haben, werden auf Grund des Prinzips der Abschattung für das im Schallschatten liegende Ohr gedämpft wahrgenommen. [PC21]

Schallwellen mit einer größeren Wellenlänge hingegen werden um das Hindernis herumgelenkt. Das Prinzip der ILD ist daher sowohl frequenzabhängig als auch richtungsabhängig. Der größte Pegelunterschied zwischen den beiden Ohren kann bei hohen Frequenzen, welche aus einem 90° Azimut einfallen beobachtet werden. Die Grenzfrequenz, bei der der Effekt der ILD nur noch in geschwächter Form auftritt liegt bei etwa 1,5 kHz. Dies ist genau jene Frequenz, auf welche im nächsten Unterpunkt näher eingegangen wird.

3.2.3 Duplex Theorie

Die von Rayleigh bereits 1907 erwähnte *Duplex Theorie* stellt die bereits genannten Effekte der ITD und der ILD in Beziehung zueinander. Er beschrieb in dieser Veröffent-

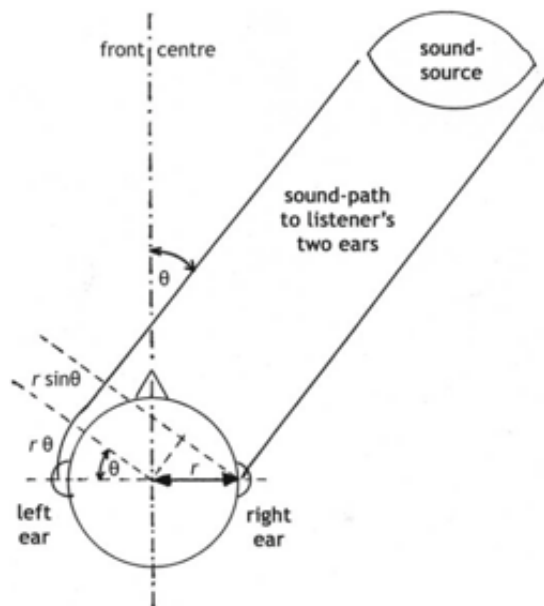


Abbildung 5 – Interaurale Laufzeitdifferent abhängig von Einfallswinkel [PC21]

lichung, dass die Fähigkeit, Klänge zu lokalisieren, sowohl von der Pegeldifferenz (Abschattung), als auch von der Laufzeitdifferenz (Wegdifferenz) abhängt, und diese beiden Effekte jeweils eine Frequenzabhängigkeit aufweisen. [Lin14]

Anders ausgedrückt besagt sie, dass zur Bestimmung des Hörereignisortes bei hohen Frequenzen die ILD und bei niedrigen Frequenzen die ITD von entscheidender Bedeutung ist. [Bra02]

Als Grenzfrequenz wird hier wie bereits erwähnt 1,5 kHz genannt, bei welcher die Schallwelle etwa eine Wellenlänge in der Größenordnung des Durchmessers des menschlichen Kopfes hat. Oberhalb dieser Frequenz kommt der Effekt der Abschattung (ILD) zum Tragen. Hier ist die ITD von geringerer Bedeutung, da sich der Zusammenhang zwischen dem Phasenunterschied der beiden, am Trommelfell eintreffenden, Ohrsignale bei zunehmend höheren Frequenzen nicht mehr korrekt interpretiert werden kann. Bei tiefen Frequenzen hingegen ist die ILD von geringerer Bedeutung, da Frequenzen mit einer großen Wellenlänge um den Kopf herumgebeugt werden. Die ITD ist hier durch den besagten Wegunterschied, den die beiden, am Trommelfell eintreffenden, Ohrsignale aufweisen, entscheidend für die Lokalisationsfähigkeit.

3.2.4 Trading Effekt

Ein interessantes psychoakustisches Phänomen, welches an dieser Stelle Erwähnung finden sollte, ist der *Trading Effekt*. Hierbei geht es um eine gegenseitige Kompensation von ITD bzw. ILD. Eine geringe Laufzeitdifferenz kann nach dieser Theorie durch eine Pegeldifferenz ausgeglichen werden, was dazu führt, dass das Hörereignis in vielen Fällen verschoben wahrgenommen wird. Diese Wechselwirkung zwischen Intensität und Laufzeit variiert je nach Studie stark und wird von $2 - 35 \mu s/dB$ angesetzt. [Moo12]

Bei zu extremen Einstellungen zerfällt das Ereignis jedoch in zwei Hörereignisse, was den Trading Effekt nur in gewissen Bereichen relevant macht. [Bra02]

Der Ursprung dieses Phänomens lässt sich auf die Reizübertragung zurückführen, und wird häufig damit erklärt, dass die benötigte Übertragungszeit der Reizübertragung für Reize mit höherer Intensität eine kürzere Zeit benötigt, als für Reize mit niedriger Intensität. Das bedeutet also, dass ein intensiverer verzögert ankommender Reiz, die selbe Reaktion auslösen kann, wie ein weniger intensiverer früher ankommender Reiz. [Moo12]

3.3 Richtungslokalisierung in der Medianebene

In der Medianebene, welche orthogonal zur Horizontalebene liegt, haben interaurale Differenzen der Ohrsignale keine Bedeutung. Da die Medianebene den Kopf in zwei symmetrische Hälften unterteilt weisen beide Ohrsignale keine ITDs bzw. ILDs auf. Wahrgenommene Klänge werden hier also für den Hörenden als *monaural* wahrgenommen. Entscheidend für eine gute Richtungswahrnehmung sind hier stattdessen spektrale Filterungseffekte, welche durch die Form des Außenohres, den Einfluss des Kopfes, sowie durch den Einfluss des Rumpfes (Torso) inklusive der Schultern entstehen.

3.3.1 Eigenschaften des Außenohres

Die primäre Aufgabe der Ohrmuschel ist es Schall der Umgebung einzufangen und die eingefangenen Schallwellen mittels dem Gehörgang an das Trommelfell zu übertragen. Durch die Ausformung der Muschel wirkt sie zusammen mit dem Gehörgang wie ein Resonanzsystem, welches das Frequenzspektrum der eintreffenden Schallwelle beeinflusst. [BB08]

Dieses Resonanzsystem ist vom Einfallswinkel des eintreffenden Schallereignisses abhängig und variiert von Mensch zu Mensch. Besonders auffallend ist jedenfalls stets die Resonanz zwischen 2 kHz - 5 kHz, welche durch die Resonanz des Gehörganges erklärt werden kann. Diese spektralen Verfärbungen alleine enthalten allerdings noch keine Information über den Einfallswinkel. Der Mensch lernt bereits in jungen Jahren diese spektrale Information mit dem Einfallswinkel von Schallereignissen zu verknüpfen und kann so aus monauralen Signalen Information über die Position von Schallquellen gewinnen.

Den größten Einfluss auf Richtungswahrnehmung übt hierbei die Breite und Tiefe des Cavum Conchae (Gehörgangseingangstrichter), sowie die Drehung des Ohres aus. Die Höhe des Ohres hingegen hat nur einen beschränkten Einfluss auf die Außenohrübertragungsfunktion sowie die wichtigsten binauralen Merkmale. Dies wird darauf zurückgeführt, dass die Ohrhöhe der Parameter ist, welcher sich im Wachstum am stärksten verändert. Durch die ständige Veränderung des Parameters, wird er für eine Interpretation der Richtungswahrnehmung beim Erlernen der auditiven Fähigkeiten, nicht als primäre Informationsquelle verwendet. [FV08]

3.3.2 Blauert'sche Bänder

Das von Jens Blauert vorgeschlagene *Modell der richtungsbestimmenden Bänder* ist bis heute anerkannt und beschreibt wieso die Richtungswahrnehmung auf der Medianebene möglich ist. Wie in Abbildung 6 zu sehen ist, lassen sich Frequenzbänder bestimmen, auf welchen gewisse Einfallsrichtungen wahrgenommen werden. So ist der Bereich um 8 kHz etwa für die Lokalisation von Schallquellen über dem Kopf verantwortlich. Der Bereich um 120 Hz bzw 1,2 kHz nimmt Schallquellen von hinten, und der Bereich um 450 Hz bzw 4,5 kHz Schallquellen von vorne wahr. Dies ergibt also ein frequenz- und richtungsabhängiges Resonanzsystem, mit welchem eine Art von Frequenzcodierung der Quellenposition durchgeführt wird. [PC21]

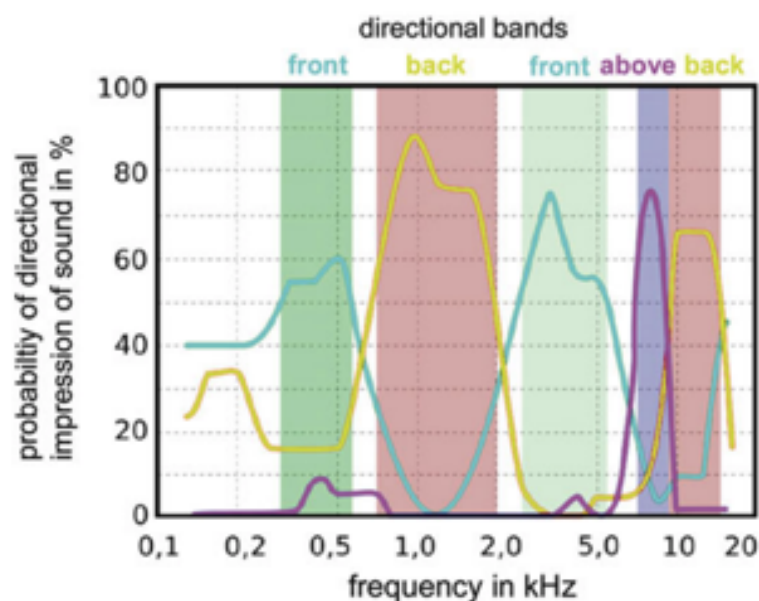


Abbildung 6 – Blauert's richtungsbezogene Bänder [PC21]

3.4 Außenohrübertragungsfunktion

All die bisher besprochenen Konzepte und Mechanismen werden in der sogenannten Außenohrübertragungsfunktion (engl.: *head related transfer function*, HRTF) vereint. Mit ihr kann der Übertragungsweg des Ursprungsortes eines Schallereignisses bis zum Trommelfell mathematisch beschrieben werden. [BB08]

Entlang dieser Übertragungsstrecke treten lineare Verzerrungen auf, welche durch die Ohrmuschel, sowie Körper und Kopf verursacht werden. Da jeder Mensch eine andere Ausprägung dieser Merkmale hat unterscheidet sich auch die HRTF bei jedem Individuum. Sie kann also quasi als eine Art "akustischer Fingerabdruck" gesehen werden. Mathematisch gesehen, kann man die HRTFs des rechten bzw. linken Ohrs wie folgt beschreiben:

$$H_R(r, \varphi, \delta, f, a) = \frac{P_R(r, \varphi, \delta, f, a)}{P_0(r, f)}$$

$$H_L(r, \varphi, \delta, f, a) = \frac{P_L(r, \varphi, \delta, f, a)}{P_0(r, f)}$$
(1)

Wobei P_R und P_L hier für die komplexwertigen Schalldrücke am rechten bzw. linken Ohr stehen und P_0 den komplexwertigen Freifeldschalldruck im Ursprung des Kopfkoordinatensystems ohne Anwesenheit des Kopfes repräsentiert. Wie schon in Abschnitt 3.1 beschrieben stehen r , φ und δ für die Abhängigkeit der Quellposition relativ zum Kopfkoordinatensystem. Die Variable f steht für die Frequenzabhängigkeit und a symbolisiert den Einfluss der menschliche Anatomie. [Xie13]

Diese anatomische Abhängigkeit ist dafür verantwortlich, dass eine Mittelung über einer große Anzahl von HRTFs oft zu keinen guten Resultaten in der Lokalisationsfähigkeit bei Hörversuchen führt, da es zwischen Proband*Innen große anatomische Unterschiede geben kann. Wird nun über diese unterschiedlichen HRTFs gemittelt entsteht eine Unschärfe, die dazu führt, dass es beim Wiedergabevorgang zu Abbildungsfehlern kommt. Es wurde in Versuchreihen gezeigt, in denen Probanden durch HRTFs anderer Personen Lokalisationstests durchgeführt haben, dass die Lokalisationsfähigkeit dadurch merklich verschlechtert wurde. [Rum01]

Für die binaurale Wiedergabetechnik stellt die HRTF jedenfalls ein wichtiges Werkzeug da, um die, am Trommelfell anliegenden, Ohrsignale zu simulieren und der hörenden Person so die Illusion zu vermitteln, dass die Klangquelle dreidimensional - wie in der Realität - zu hören ist. Diesen Vorgang nennt man die *Auralisation*. [BB08]

3.5 Entfernungswahrnehmung

Die Schallintensität bietet einen wichtigen Anhaltspunkt um Entfernungen akustisch einschätzen zu können. Diese wird durch die Mechanismen des menschlichen Gehörs in eine frequenzabhängige Lautheitsempfindung umgewandelt. Grundsätzlich gilt unter Freifeldbedingungen für Punktschallquellen eine 6 dB Abnahme des Schalldrucks je Entfernungsverdoppelung. Für eine Halbierung der Lautheit ist dies aber nicht so einfach umzurechnen, da Lautheit in Sone gemessen wird. Man kann allerdings näherungsweise abschätzen, dass eine Halbierung der Lautheitsempfindung für Frequenzen von 400 Hz bis 5 kHz, ab einem Schalldruckpegel von 40 dB und je nach Signal, in etwa mit einer Schalldruckabnahme von 10 dB korrespondiert.

Die bereits genannte Abnahme des Schalldrucks ist allerdings nur unter Freifeldbedingungen gültig. Sobald sich die Hörer im diffusen Schallfeld befindet ändert sich diese Abnahme, sowie die damit verbundene wahrgenommene Lautheit des Hörereignisses. Da die Schallenergie im diffusen Schallfeld deutlich langsamer abnimmt ist die Wahrnehmung und Einschätzung von Entfernungen eine Andere. Das Verhältnis von Direktschall zu Diffusschall ist daher für die Entfernungseinschätzung in Räumen ein noch wichtigerer

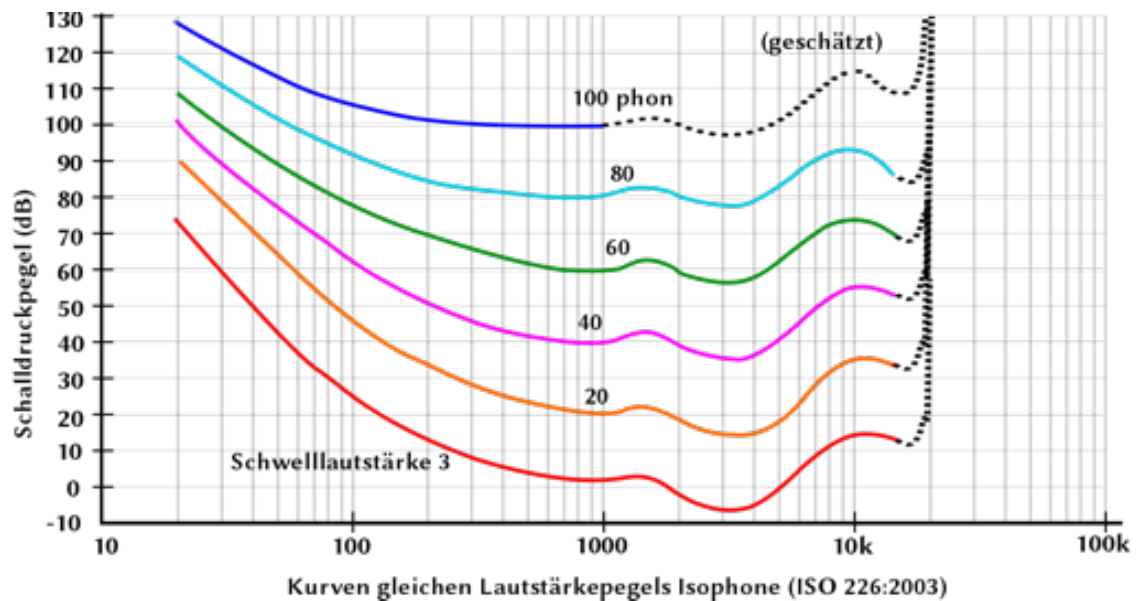


Abbildung 7 – Kurve der gehörrichtigen Lautstärke

Indikator als der Schalldruckpegel. Im Speziellen ist das Verhältnis von Direktschallenergie zu Energie der frühen Reflexionen ($< 80ms$) entscheidend.

Es fließt in die Entfernungsabschätzung aber auch die Frage ein, ob der Hörende mit seiner Umgebung vertraut ist. Wenn die klanglichen Eigenschaften des Raumes nicht bekannt sind, ist es deutlich schwieriger die Entfernung einzuschätzen. Erklärt wird dieses Phänomen durch die sogenannte assoziative Mustererkennung, da der menschliche Hörmechanismus in der Lage ist die akustische Information eines bekannten Raumes zu extrahieren.

Die Luft verursacht bei hohen Frequenzen eine Dämpfung der Schallausbreitung (*Dissipation*), was dazu führt, dass auch die Entfernung selbst eine Auswirkung auf das Frequenzspektrum hat. Der Mensch erlernt durch die Interpretation dieser spektralen "Hinweise" bereits in jungen Jahren wie Entfernungen einzuschätzen sind. [PC21]

Allgemein gilt, dass die menschliche Entfernungseinschätzung nicht sonderlich exakt ist und es selbst bei bereits bekannten Klangquellen, in vertrauten Räumen häufig zu Fehleinschätzungen von etwa 20 % kommen kann. [Moo12]

3.6 Lokalisation von mehreren Schallquellen

Bisher wurden lediglich Betrachtungen angestellt, bei denen nur eine Schallquelle involviert war. Ergibt sich allerdings eine *Mehrschallquellsituation*, so ist es dem Menschen nicht immer möglich den Einzelschallquellen eine exakte Richtung zuzuordnen. Diese im Raum klar getrennten Schallereignisse beeinflussen sich bei der Wahrnehmung gegenseitig, sodass die wahrgenommenen Hörereignisse nicht immer getrennt als solche identifiziert werden können. In diesem Kapitel werden einige dieser psychoakustischen Phänomene besprochen. [BB08]

Präzedenzeffekt Der auch unter dem Namen des *Gesetzes der ersten Wellenfront* bekannte *Präzedenzeffekt*, ist dafür verantwortlich, dass der Mensch in geschlossenen Räumen die Ur-Schallquelle ausmachen kann und sich nicht von den Reflexionen irritieren lässt. Man ist so in der Lage die Richtung der ersten Wellenfront auszumachen, da die frühen Reflexionen erst nach einer zeitlichen Verzögerung ankommen. Überschreitet diese Verzögerung jedoch einen Grenzwert (die *Echoschwelle*), so nimmt man nicht mehr ein, sondern zwei Schallereignisse wahr. Diese Echoschwelle liegt in der Praxis je nach Signalcharakteristik bei etwa 5 bis 100 ms. [BB08]

Summenlokalisation Strahlen zwei oder mehrere Schallquellen ein sehr ähnliches, oder identes Signal ab, so entsteht ein Hörereignis, dessen Ort von den Orten und den Signalen der verschiedenen Schallquellen abhängt. Dieses Phänomen nennt man die Summenlokalisation. [Wei08]

Wenn beispielsweise von zwei Lautsprechern, welche in exakt gleicher Entfernung zum Zuhörer positioniert sind, das exakt gleiche Signal abgestrahlt wird, so sorgt die Summenlokalisation dafür, dass der Mensch nur eine fiktive Schallquelle wahrnimmt - eine sogenannte *Phantomschallquelle*. Die wahrgenommene Schallquelle wird allerdings weniger scharf lokalisiert, als eine Einzelschallquelle mit demselben Schalldruck. [BB08]

Assoziationsmodell Ein alternatives Erklärungsmodell für die Lokalisationsfähigkeit ist das Modell der *assoziativen Mustererkennung*. Dabei geht es darum, dass im Gehirn bei der Reizverarbeitung aktuelle Muster mit bereits gespeicherten Mustern verglichen werden. Dies geschieht auch dann, wenn nur Teile des Musters korrelieren. Auf diese Weise kann der Mensch beispielsweise Rückschlüsse über räumliche Information ziehen, bzw. umgekehrt auch in einem bekannten Raum über das unbekannte abgespielte Signal. [The80]

4 Binaurale Aufnahmesysteme

4.1 Prinzip von binauralem Stereo

Wie schon in Kapitel 2 erwähnt, wird das Prinzip der binauralen Aufnahme- bzw. Wiedergabetechnik seit vielen Jahren erforscht. Doch was ist das grundsätzliche Prinzip dieser Technologie und warum sollte binaurales Audio keinesfalls mit herkömmlichen Stereoaufnahmen über einen Kamm geschert werden?

Prinzipiell stimmt es, dass auch die binauralen Aufnahmesysteme zur Gruppe der 2-Kanal-Stereoaufnahmeverfahren gezählt werden können. Jedoch unterscheidet man hier zwischen der sogenannten Lautsprecherstereophonie und der kopfbezogenen Stereophonie. Letztere ist für die Aufnahme von binauralen Audiosignalen verantwortlich. Ziel dieses Verfahrens ist es, dass das aufgenommene Signal einer Person bei der Wiedergabe den Höreindruck vermittelt, der in einer realen Hörsituation wahrgenommen worden wäre. [RM21]

Hierzu werden auf zwei Kanälen die Ohrsignale aufgezeichnet, welche sowohl Informationen über Klangquelle inklusive Umgebung, als auch Informationen über spektrale Verzerrungen enthalten, welche sich auf Grund des Einflusses der menschlichen Anatomie auf die Schallereignisse ergeben. Dies beinhaltet auch das Konzept der HRTF, welche ausführlich in Kapitel 3.4 behandelt wurde. Eine wichtige Berücksichtigung für den Themenschwerpunkt des binauralen Audios sind allerdings nicht nur die Aufnahme-, sondern auch die Wiedergabemethoden. Ein binaurales System sollte stets als Gesamtsystem - die gesamte Aufnahme- und Wiedergabekette - betrachtet werden, da nur so das Ziel der originalgetreuen Reproduktion eines Klangereignisses erreicht werden kann.

Da die aufgenommenen Ohrsignale den Schalldruck im Innenohr repräsentieren, sollte eine Wiedergabe im Normalfall über Kopfhörer stattfinden, da nur so die nötige Kanaltrennung gewährleistet werden kann. Verfahren, die diese Kanaltrennung bei Lautsprecherwiedergabe mit Hilfe der sogenannten "crosstalk cancellation" realisieren, sind zumeist nur unter großem Aufwand umsetzbar. Dem kommt außerdem hinzu, dass sowohl bei einer Wiedergabe über Lautsprecher, als auch zum Teil über Kopfhörer, die, der menschlichen Anatomie geschuldeten, spektralen Verfärbungen den Hörenden doppelt erreichen würden, da diese schon in der Aufnahme enthalten sind. [WBGC⁺17]

Eine solche binaurale Aufnahme kann im Allgemeinen nun entweder über einen Kunstkopf oder über einen Echkopf erfolgen. Auf diese beiden Methoden, sowie die mit ihnen verbundenen Anforderungen, Qualitätskriterien, sowie Problemstellungen wird in weiterer Folge in diesem Kapitel näher eingegangen.

4.2 Aufbau des Kunstkopfes

Ein Kunstkopf ist eine artifizielle Nachbildung der anatomischen Strukturen des Menschen, welche die Dimensionen einer durchschnittlichen Zuhörer*in hat. [Vor04]

Dies erfolgt über eine Nachbildung des Kopfes, des Außenohres, sowie des Oberkörpers. Es sollte allerdings angemerkt werden, dass die Ausprägung der jeweiligen Merkmale von Modell zu Modell variieren. Die Art der Anwendung ist ein wichtiger Punkt bei der

Beschreibung des Aufbaus von Kunstköpfen, da hier je nach Funktion Unterschiede bestehen an welcher Stelle die, für die Aufnahme nötigen, Mikrofone befestigt werden. Prinzipiell differenziert man bei Kunstkopfsystemen zwischen Kunstköpfen für ohrferne bzw. ohrnahe Schallquellen. Diese Unterscheidung ist maßgeblich, ob im System der Ohrkanal sowie Trommelfellimpedanz nachgebildet wird. Bei der Messung von ohrnahen Schallquellen - wie zum Beispiel bei der Ausmessung von Kopfhörern - hat der Ohrkanal maßgeblichen Einfluss auf die akustische Last der Quelle, sowie die damit verbundene abgestrahlte Leistung. Hier muss Ohrkanal und Trommelfellimpedanz nachgebildet werden. Bei der Messung von ohrfernen Schallquellen kann dies hingegen vernachlässigt werden. [DFF⁺07]

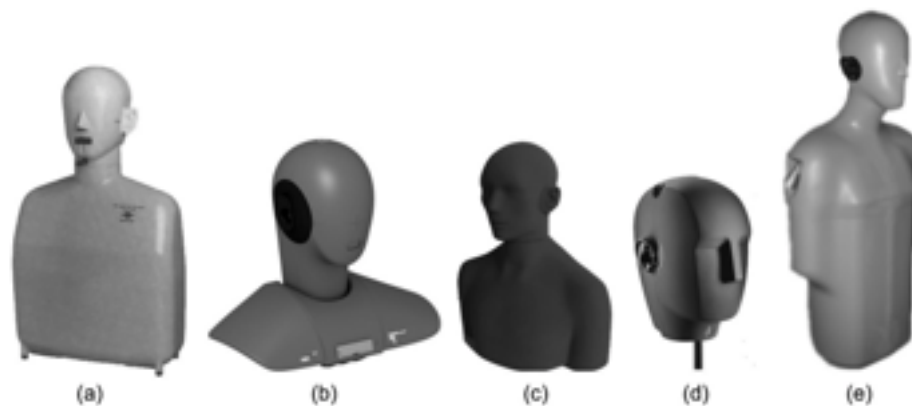
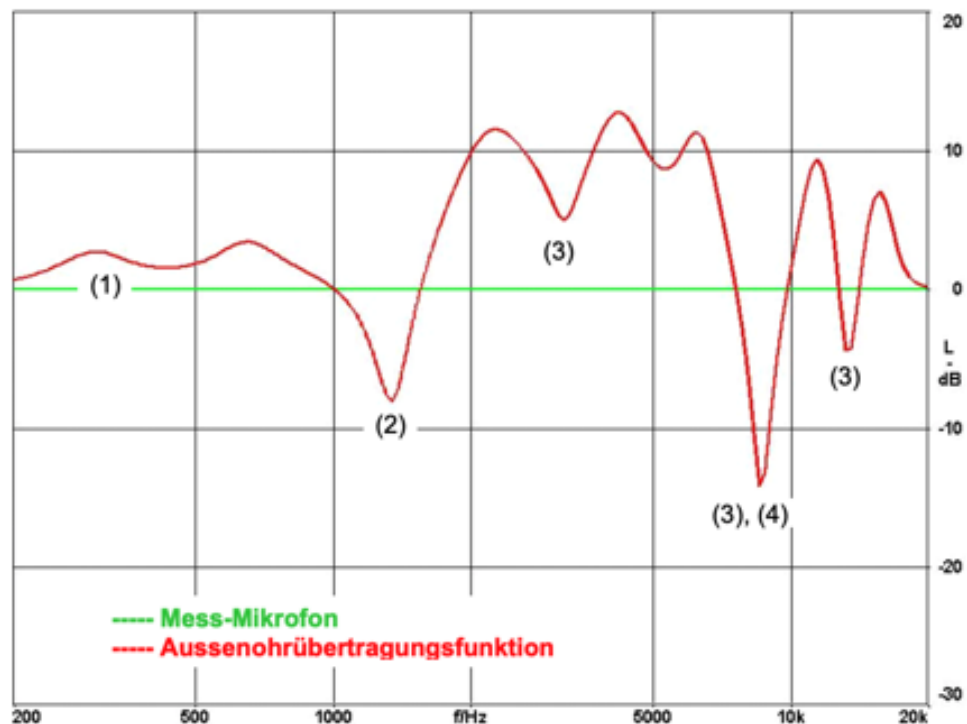


Abbildung 8 – Verschiedenen Kunstkopfmodelle im Überblick: (a) Brüel & Kjør 4128; (b) HEAD acoustics HMS IV; (c) 01dB-Metravib MK2B; (d) Neumann KU-100; (e) VALDEMAR [Xie13]

Bei der Form des Kopfes spielt vor allem der gewählte Abstand der beiden Ohren eine entscheidende Rolle, da sie die resultierende ITD der beiden Mikrofone beeinflusst und mit dem Ohrenabstand der hörenden Person übereinstimmen muss. Ein weiterer wichtiger Aspekt ist die originalgetreue Rekonstruktion des Außenohres, welches zumeist aus elastischen Materialien wie Silikon hergestellt wird. Die Form und vor allem die Größe der Ohren ist dabei entscheidend, da sie für die Lokalisationsfähigkeit essentiell sind. Je nach Modell wird häufig auch der Oberkörper (Torso) nachgebaut, da dadurch der Einfluss des Oberkörpers auf das Hörereignis nachgebildet werden kann.

In welcher Form sich die beschriebenen Merkmale auf die Außenohrübertragungsfunktion auswirken ist in Abbildung 4.2 dargestellt. (1) zeigt den tieffrequenten Bereich, welcher bei einer Wellenlänge $\lambda \gg$ Abmessung des Kopfes auftritt. Hier verhält sich die HRTF annähernd wie der Frequenzgang des linearen Messmikrofons. Der Bereich (2) zeigt destruktive Interferenzen, die sich durch die Schulterreflexionen ergeben. (4) zeigt den Einfluss des Außenohres auf den Frequenzgang und unter (3) kann man erkennen, dass auch die $\lambda/4$ Resonanz des Ohrkanals entscheidend auf die HRTF einwirkt. [DFF⁺07]

Diese Messung wurde bei einer realen Person durchgeführt, verdeutlicht jedoch gut, warum



- (1) Wellenlänge $\lambda \gg$ Abmessungen des Kopfes
- (2) Einfluss der Schulter
- (3) Einfluss des Ohrkanals ($\lambda/4$ -Resonanz)
- (4) Einfluss des Außenohres

Abbildung 9 – Einfluss der menschlichen Anatomie auf die HRTF einer Versuchsperson bei frontaler Beschallung am linken Ohr [DFE⁺07]

es nötig ist, bei der Konstruktion eines Kunstkopfes derart hohen Aufwand zu treiben. Bei einem Modell ohne Torso würde bei einer Messung der Bereich (2) beispielsweise näherungsweise einen flachen Verlauf aufweisen.

4.3 Echkopfmikrofonierung als Alternative zum Kunstkopf

Im Gegensatz zur Aufnahme mit Hilfe eines Kunstkopfsystems ist eine alternative Lösung hierzu die Verwendung einer realen Person um Schallereignisse aufzuzeichnen. In diesem Fall werden einer Proband*in zwei Kopfmikrofone (Druckempfängermikrofone) in die Eingänge der Gehörgänge gesteckt.

Ein Vorteil dieser Variante ist jedenfalls die einfache Realisierung, sowie die Verwendung der eignen Ohren. Wie schon in Unterkapitel 3.4 besprochen wurde, können auf diese Weise die Hörereignisse aus der Sicht der Versuchsperson am besten abgebildet werden. Es wurde in einigen Studien sogar gezeigt, dass mit einer gut ausgewählten Versuchsperson bessere Resultate im Bezug auf die Lokalisationsfähigkeit erreicht werden konnten,

als mit guten Kopfkopfsystemen. Im Detail wird darauf in Unterkapitel 4.8 eingegangen. Den Vorteilen dieser günstigeren Alternative stehen jedoch auch Nachteile gegenüber. Diese zeigen sich, in Form eines niedrigen Signal zu Rausch Verhältnisses (engl. signal to noise ratio, SNR) des Signals durch die meist sehr klein gehaltenen Wandler in den Kopfmikrofonen, die Nicht-Kompatibilität zu anderen Personen, welche eine stark abweichende HRTF zu jener der Versuchsperson aufweisen, sowie eine gewisse Unschärfe in der Aufnahme, welche den unvermeidbaren Kopfbewegungen der Proband*in geschuldet ist. [Wei08]



Abbildung 10 – Kopfmikrofonanfertigung - realisiert mit Hilfe eines Sennheiser KE 4-211-2 Mikrofons am Eingang des Ohrkanals einer realen Person - angefertigt an der Universität Aalborg [HM05]

4.4 Prinzip der Aufnahme- und Wiedergabekette

Um bei Hörtests eine qualitativ hochwertige Aussage treffen zu können, sollte der Aufnahmeprozess niemals isoliert betrachtet werden. Die Wechselwirkung mit dem Wiedergabemedium - Kopfhörer oder Lautsprecher - sowie die Einflüsse der Übertragungskette müssen stets in die Bewertung der Aufnahme mit einfließen. Das aufgenommene binaurale Signal wird im Gehörgang von Mikrofonen aufgezeichnet und repräsentiert den Schalldruck, den eine Person im Originalschallfeld wahrgenommen hätte. Der Schalldruck, welcher bei der Reproduktion erzeugt wird, muss nun exakt mit dem Schalldruck übereinstimmen, der bei der Aufnahme im Originalschallfeld geherrscht hat. Nur so kann der hörenden Person ein möglichst realer räumlicher Gesamteindruck vermittelt werden. Im Idealfall wird das binaurale Signal mit den eignen Ohren der hörenden Person aufgenommen, was dazu führt, dass bei der Reproduktion des Signals eine authentische Wiedergabe des Originalsignals gewährleistet werden kann. Diese Form der Aufnahme-Wiedergabekette kann jedoch in vielen Fällen nicht umgesetzt werden, oder hat sich für manche Anwendungen als zu umständlich herausgestellt. Eine Aufnahme erfolgt daher in den meisten Fällen entweder mit Hilfe eines Kopfkopfes, oder durch eine ausgewählten Person, deren Außenohrübertragungsfunktion eine gute Annäherung an die Au-

Benohrübertragungsfunktionen möglichst vieler Personen darstellt. In beiden dieser Fälle ist eine originalgetreue Reproduktion ohne die Durchführung einiger signalverarbeitungs-theoretischer Schritte nicht möglich. Schon kleine Abweichungen in der Außenohrübertragungs-funktion können zu Unterschieden bei den reproduzierten Schalldrücken am Trom-melfell der hörenden Person führen. Die Ungleichheit in den Frequenzgängen von Auf-nahmemikrofon zu Wiedergabemedium führt zu linearen Verzerrungen im Amplituden-und Phasengang des reproduzierten Schalldrucks. Um diese Verzerrungen zu umgehen bzw. zu verringern kann in der Übertragungskette ein Equalizer eingesetzt werden. Dieser besteht für die beiden Ohrsignale aus der Inversen der nichtlinearen Übertragungsfunktion $H_{pL}(f)$ und $H_{pR}(f)$, und berechnet sich wie folgt :

$$\begin{aligned} F_L(f) &= \frac{1}{H_{pL}(f)} \\ F_R(f) &= \frac{1}{H_{pR}(f)} \end{aligned} \quad (2)$$

Diese Korrektur der Übertragungsfunktion wird vor der Wiedergabe auf das Signal ange-wendet und ermöglicht so, dass dem Hörenden ein besserer räumlicher Gesamteindruck vermittelt werden kann.

Abbildung 11 zeigt diese schematische Darstellung anhand eines Blockdiagramms. [Xie13]

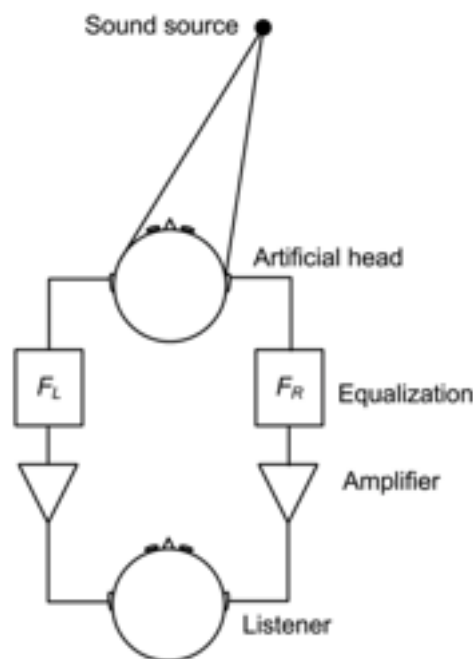


Abbildung 11 – Blockdiagramm der binauralen Aufnahme-Wiedergabekette [Xie13]

In dieser Betrachtung ist nicht auf Aspekte wie die korrekte A/D bzw. D/A Wandlung der Audiosignale eingegangen worden. Auch diese kann bei der Reproduktion zu Unschärfe des Hörereignisses führen und sollte nicht außer Acht gelassen werden. Ebenfalls wurde

nicht auf die Problematik der Wiedergabe über Lautsprecher eingegangen. Diese wird in Unterpunkt 4.6 näher behandelt.

Kompatibilität mit Lausprechern Die Wiedergabe von binauralen Signalen über Lautsprecher wird *transaurale Wiedergabe* genannt. Hierbei wird das aufgenommene Signal anstelle von Kopfhörern, über Lautsprecher wiedergegeben, wodurch sich mehrere Problemstellungen ergeben:

1. Es fehlt bei der Wiedergabe die Kanaltrennung. Das für das linke Ohr bestimmte Signal erreicht das linke Ohr und etwas zeitverzögert bzw. spektral verändert das rechte Ohr. Vice versa kommt es zu diesem Effekt natürlich auch beim zweiten Ohr.
2. Die Signale enthalten vor der Wiedergabe bereits alle binauralen Merkmale. Durch die Wiedergabe wird die binaurale Übertragungsstrecke nun noch einmal durchlaufen.
3. Raumakustische Eigenschaften des Wiedergaberaumes haben Einfluss auf das gehörte Signal und haben somit negativen Einfluss auf die Qualität des wiedergegebenen Signals

Diese drei Effekte führen zu einer Verschlechterung der Abbildungsschärfe, was dazu führt, dass auch der räumlichen Eindrucks darunter leidet. Bei nachhallbehafteten Wiedergaberräumen können mit Hilfe einer Raumentzerrung Teile des Problems behoben werden, jedoch kommt es auch so zu Qualitätsverlusten im Wiedergabeprozess. [SBPW16]

Für das Minimieren des sogenannten *Interaural Crosstalks* wurde die *Crosstalk Cancellation* (CTC) entwickelt, durch welche ein Großteil des Übersprechproblems zwischen den Lautsprechern eliminiert werden kann. Allerdings ist auch diese Technologie nur eine Kompromisslösung, da die nötige Kanaltrennung nur in einem sehr kleinen Sweet Spot - der idealen Hörposition - erreicht werden kann.

Der zweite angeführte Punkt kann durch eine inverse Filterung des Frequenzspektrums der binauralen Übertragungsstrecke minimiert werden und wirkt so gegen die dadurch ausgelösten spektralen Verfärbungen. [Wei08]

4.5 Anwendungsgebiete

Die binaurale Aufnahmetechnik hat sich in einer Vielzahl von Einsatzgebieten über die Jahre als sehr hilfreich erwiesen, da sie es ermöglicht, die auditive Wahrnehmung des Menschen besser zu verstehen.

So hat sich etwa in der Psychoakustik bei diversen Hörtests der Kunstkopf als eine verlässliche *Referenzversuchsperson* als nützlich erwiesen. Hierfür werden großteils Kunstköpfe verwendet, da sie lineare, zeitinvariante und in allen Testsituationen überprüfbare Übertragungseigenschaften aufweisen. Mit Hilfe dieser Technik kann nicht nur der Schall- sondern das Hörereignis, welches die kognitiven Aspekte der Wahrnehmung des Menschen berücksichtigt, betrachtet werden.

Ein wichtiger Anwendungsbereich liegt beispielsweise in der Automobilindustrie, um die *Fahrzeuginnenraumakustik* zu analysieren. Auch die Untersuchung von Umwelt- und Arbeitsplatzgeräuschbelastung, welche unter psychoakustischen Gesichtspunkten betrachtet werden, fällt in das Anwendungsgebiet der binauralen Aufnahmetechnik. Im Bereich der *Vermessung von Hörgeräten* wird der Kunstkopf für die Untersuchung von Algorithmen zur Unterdrückung von Hintergrundgeräuschen eingesetzt, da er eine Möglichkeit bietet, den Einfluss der Rauschunterdrückung auf die Sprachverständlichkeit zu testen. [DFF⁺07]

Eine weitere Anwendung ist die *Lautsprecher-simulation*. Hierbei wird die binaurale Raumimpulsantwort von Lautsprechern gemessen um den Klangeindruck im Anschluss daran für beliebige Quellsignale binaural zu auralisieren. Dies ist für die subjektive Bewertung verschiedener Lautsprechermodelle interessant, da der raumakustische Einfluss damit verringert werden kann. [MG06]

Auch im *architektonischen Planungsprozess* kommt die binaurale Messtechnik zum Einsatz. In diesem Zusammenhang wird sie für den Entwurf von Räumen unter dem Einsatz binauraler Raumsimulationstechnik verwendet, und anschließend über Kopfhörer via Headtracking auralisiert. Der Kunstkopf dient in dieser Hinsicht zur Verifikation und Datenakquise. [Mol03]

Die Aufnahmetechnik ist ein weiterer Bereich in der der Kunstkopf eingesetzt wird. Hier bietet sich etwa die Möglichkeit die binaurale Aufnahmetechnik für *Vergleiche von Konzertsälen* einzusetzen. Durch binaurale Aufnahmen kann beispielsweise der akustische Eindruck des selben Orchesters in unterschiedlichen Konzertsälen ohne Zeitverzögerung verglichen werden. [Møl92]

Ein Weiterer Aspekt liegt hier im Einsatz des Kunstkopfes als Hauptmikrofon. Dies wird jedoch auf Grund der schlechten Kompatibilität zu Lautsprechersystemen häufig widersprüchlich diskutiert. Dieses Kapitel bietet jedenfalls nur einen kleinen Ausschnitt in das umfangreiche Einsatzgebiet der binauralen Aufnahmetechnik, welches laufend durch neue Anwendungen erweitert wird.

4.6 Herausforderungen bei binauralen Systemen

Bis heute konnte die Entwicklung eines optimalen binauralen Aufnahmesystems, welches es ermöglichen würde abbildungsfehlerfreie binaurale Aufnahmen vorzunehmen, nicht realisiert werden. Zu groß sind die technischen Hindernisse eines solchen Systems, die hierfür überwunden werden müssten. Auf einige dieser Problemstellungen wurde in dieser Arbeit schon aufmerksam gemacht. Es folgt nun eine Betrachtung von einigen dieser problematischen Aspekte.

Individualproblem Bei allen binauralen Systemen muss stets zwischen Kompatibilität mit möglichst vielen Versuchspersonen sowie dem möglichst realitätsgetreuen Höreindruck abgewogen werden. Beide Aspekte schließen sich in ihren Idealformen gegenseitig aus, weshalb es der Konstrukteur*in des Kunstkopfes obliegt einen guten Kompromiss zu finden. Da die Abmessungen der Kopfgeometrie einen starken Einfluss auf die Außenohrübertragungsfunktion haben, sollten sie mit höchster Sorgfalt gewählt werden. Be-

sonders zu hohen Frequenzen hin, sind unter Versuchspersonen starke Unterschiede im Frequenzgang der HRTF zu beobachten. Eine Darstellung der Standardabweichung der gemessenen HRTFs von 12 Versuchspersonen ist in Abbildung 12 zu sehen.

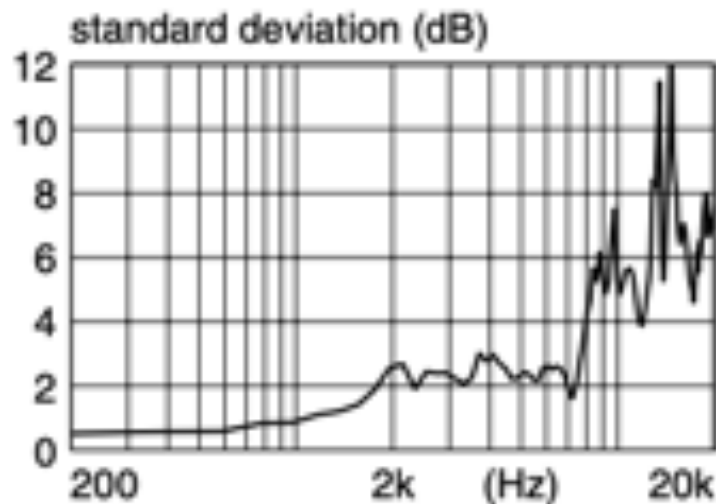


Abbildung 12 – Standardabweichung der HRTF von 12 Versuchspersonen, gemessen am Trommelfell [HM05]

Klar ersichtlich ist hier jedenfalls, dass oberhalb von 10 kHz die Abweichungen zwischen den Versuchspersonen bis zu 12 dB betragen, was dazu führt, dass erhebliche räumliche Abbildungsfehler die Folge sind.

Die Kunstkopftechnik versucht daher ein geeignetes Modell zu finden, das für möglichst viele Versuchspersonen eine gute räumliche Abbildung gewährleistet. [DFF⁺07]

Untersuchungen, auf die in Abschnitt 4.8 näher eingegangen wird, haben jedenfalls ergeben, dass mit Messungen einer gut ausgewählte Versuchsperson, welche für möglichst viele hörende Personen nur eine geringe Abbildungsunschärfe erzeugt, bessere Messergebnisse erzielt werden können, als mit den derzeit qualitativ hochwertigsten Kunstköpfen. [Vor04]

Fehlende Interaktivität Ein weiterer problematischer Aspekt, der bei der statischen binauralen Aufnahme mit Hilfe eines Kunstkopfes nicht berücksichtigt werden kann, ist die fehlende Interaktivität. Wenn ein aufgenommenes binaurales Signal wiedergegeben wird, ändert sich das Signal nicht wenn die hörende Person seinen Kopf bewegt. Dies führt zu einem unnatürlichen Hörerlebnis, da der Mensch die Bewegung des Kopfes bei der Richtungswahrnehmung zur Ortung von Klangquellen nutzt. [WBG⁺17]

Durch die fehlende Interaktivität leidet also auch die Richtungswahrnehmungsfähigkeit des Menschen. Eine Alternative hierzu stellt die in Abbildung 13 schematisch dargestellte Motion-Tracked Binaural (MTB) Technologie dar, welche eine Möglichkeit der dynamischen binauralen Aufnahme bietet. Hierzu wird mit Hilfe von einem auf einer Kugel angebrachtem Mikrofonarray das Schallfeld um die Kugel eingefangen. Dabei ist der Radius der Kugel in etwa gleich dem durchschnittlichen Abstand zwischen den Ohren eines

Menschen. Anschließend wird die Head-Tracking-Technologie eingesetzt um die Orientierung der Kopfposition der hörenden Person zu erkennen. Dadurch ist es möglich der Proband*in immer nur jene zwei Audiosignale vorzuspielen, welche mit der derzeitigen Kopfposition übereinstimmen. [Xie13]

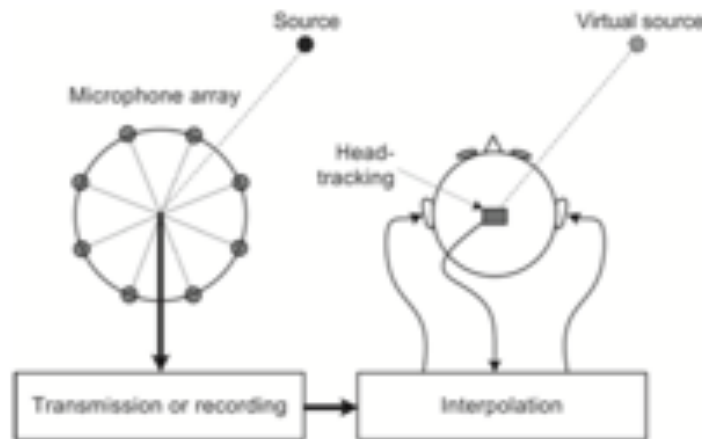


Abbildung 13 – Schema der Motion-tracked Binaural Technologie [Xie13]

4.7 Anforderungen an ein hochwertiges Aufnahmesystem

Die bisher besprochenen Methoden für die Aufnahme von binauralen Signalen bringen unterschiedliche Vor- bzw. Nachteile mit sich. Bis heute gibt es noch keine optimale Lösung um derartige Aufnahmen zu tätigen. In weiterer Folge wird nun auf die Anforderungen eingegangen, welche an ein hochwertiges Aufnahmesystem gestellt werden.

Die wichtigste Anforderung stellt die *Richtcharakteristik* dar. Der Mensch ist auf diese, wie in Kapitel 3 ausgeführt wurde, besonders sensibel und daher ist eine korrekte Richtungsabbildung von hoher Priorität. Bereits minimale Änderungen von etwa 1° können vom Menschen als solche wahrgenommen werden. Das *Eigenrauschen* sollte möglichst gering gehalten werden und unterhalb der Hörschwelle des Menschen liegen. So kommt es zu keinen Dynamikverlusten und es können auch Hörversuche im Bereich der Hörschwelle durchgeführt werden.

Auf die *Dynamik* sollte selbstverständlich auch hinsichtlich Signalen mit hoher Schallintensität geachtet werden. So sollte diese im Idealfall die Schmerzgrenze des menschlichen Gehörs überschreiten um etwaige Verzerrungen bei Pegelspitzen zu vermeiden.

Weiters spielt auch die *Kalibrierfähigkeit* eines Systems eine wichtige Rolle um mögliche systematische Fehler bei Hörversuchen ausschließen zu können.

Wie schon in Unterkapitel 4.2 erwähnt gibt es für unterschiedliche Aufnahmesituationen verschiedene Systeme, die zum Einsatz kommen. Sollen nun Schallquellen in der Nähe des Ohres aufgenommen werden, so muss darauf geachtet werden, dass das System über eine *Impedanzanpassung* verfügt. [DFF⁺07]

Auch im Bezug auf die *Reproduzierbarkeit der Aufnahme- und Wiedergabesituation* werden gewisse Anforderung an das binaurale System gestellt. Es wurde schon darüber be-

richtet, dass eine sorgfältig ausgewählte Versuchsperson, welche für möglichst viele Personen eine geringe Abbildungsunschärfe erzeugt, zuverlässigere Messergebnisse im Bezug auf Abbildungsfehler erzeugt als ein Kunstkopf. Doch trotz dieser Tatsache hat sich der Messvorgang mit Hilfe des Kunstkopfes als in der Praxis nützlicher erwiesen. Dies beruht auf der Tatsache, dass Aufnahmen mit einer sorgfältig ausgewählten Versuchsperson an einer schlechten Reproduzierbarkeit der Messungen leiden. Sowohl organisatorisch, als auch in der Messung selbst ist es schwierig mit einer realen Person mehrmals die selben Messergebnisse zu reproduzieren, da diese selbst durch minimale Kopfbewegungen verfälscht werden und es in organisatorischer Hinsicht nicht immer möglich sein wird eine terminliche Vereinbarung mit der Person zu finden.

Die schlechte Reproduzierbarkeit der Aufnahmesituation ist vor allem für die industrielle Messtechnik ein Problem, da bei fehlender Standardisierung von Kunstkopfsystemen verschiedener Hersteller der Vergleich unter binauralen Messungen oft schwierig ist. [WBG⁺17]

Auch bei der Aufnahme mit dem selben Kunstkopf in der selben Hörumgebung gibt es in dieser Hinsicht Probleme. Bei der Wiedergabe der aufgenommenen Signale gibt es je nach exakter Positionierung der Kopfhörer auf dem Kopf ebenfalls deutliche Unterschiede, welche ebenfalls eine Abbildungsunschärfe erzeugen. Es wurde gezeigt, dass durch minimale Repositionierung der Kopfhörer eine messtechnisch signifikante Verfärbungen im Frequenzspektrum bei hohen Frequenzen die Folge ist. [WBG⁺17]

4.8 Vergleich von Messungen realer Systeme

In diesem Kapitel folgt nun eine Übersicht der Ergebnisse diverser Studien, welche die Qualität von binauralen Aufnahmesystemen untersucht haben. Das Lokalisationsvermögen für unterschiedliche Systeme wurde untersucht bzw. bei unterschiedlicher Parametrisierung getestet, und daraus wurden Schlussfolgerungen über die Qualität der Aufnahmesysteme getroffen. Man spielte diesen Testpersonen eine Kunstkopfaufnahme der Testsignale, eine Echkopfaufnahme der Testsignale, sowie die Testsignale direkt über Lautsprecher (aus realer Quellrichtung) vor. Das direkt vorgespielte Testsignal diente hierbei als Referenzwert. Die Proband*innen mussten die Richtung des Signals erkennen und den für sie wahrgenommenen Lautsprecher notieren. Je weniger Fehler bei der Lokalisation der Schallereignisse gemacht wurden, umso besser wurde das System eingestuft. Zusätzlich wurde überprüft welche Art von Lokalisationsfehler gemacht wurden. Hierbei wurden Fehler in "out-of-cone error", "within-cone error", "median-plane error", sowie "distance error" unterteilt. [HM05]

Beim "within-cone error" handelt es sich um eine Vorne-Hinten-Vertauschung entlang des "cone of confusion", auf dem die Einfallsrichtungen mit der gleichen interauralen Laufzeit- und Pegeldifferenz liegen. [Wei08]

Zu einem "out-of-cone error" kam es dann, wenn es sich um einen groben Richtungsfehler handelte, der außerhalb des "cone of confusion" lag. Um einen "median-plane error" handelte es sich wenn die Richtung nicht korrekt war, aber der Fehler auf der Medianebene lag, und zu einem "distance error" kam es, wenn die Richtung des Signals stimmte, jedoch

die Distanz falsch eingeschätzt wurde. [HM05]

Eine Studie an der Universität Aalborg in Dänemark zeigte, dass die Lokalisationsfähigkeit mit Hilfe der Echkopfmikrofonierung im Durchschnitt deutlich besser, als bei Aufnahmen mit Kunstköpfen war.

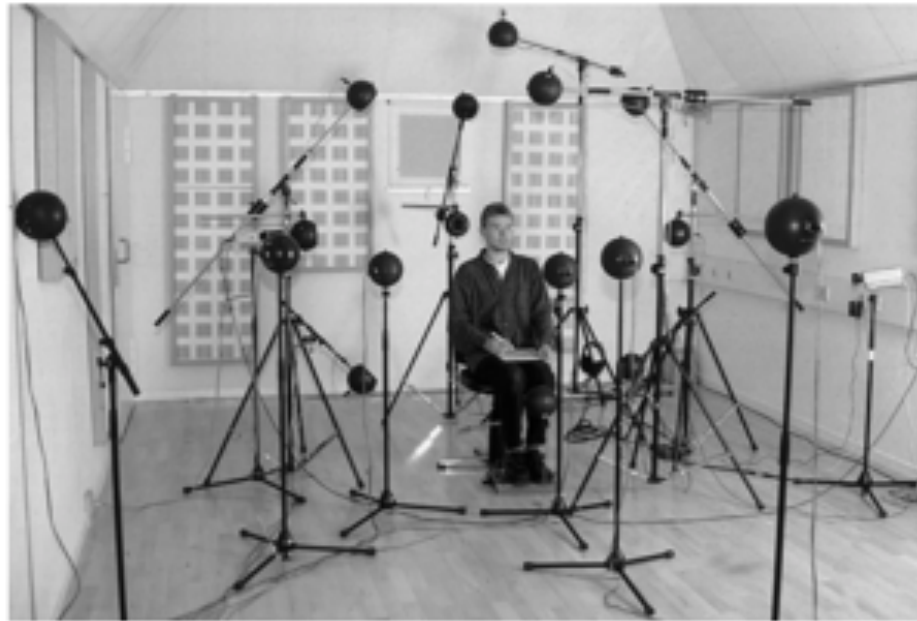


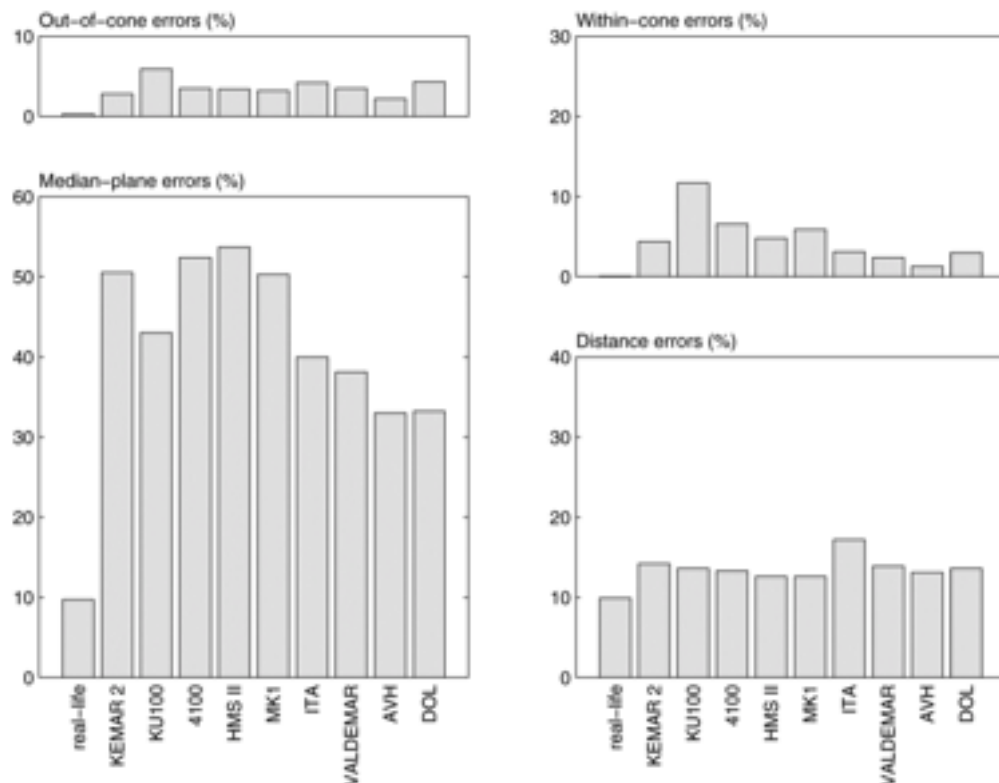
Abbildung 14 – Testumgebung eines Hörtests [HM05]

Es konnte zudem beobachtet werden, dass die Unterschiede in der Lokalisationsfähigkeit unter Menschen geringer war, als zwischen Echkopf und Kunstkopf. Daraus konnte geschlussfolgert werden, dass die Qualität eines Kunstkopfsystems umso besser ist, je genauer die geometrischen Strukturen einer realen Person nachempfunden sind. [HM05]

Eine weitere Studie an der selben Universität ist zum Ergebnis gekommen, dass Aufnahmen, welche mit fremden Ohren gemacht wurden zu deutlich mehr Lokalisationsfehlern führten, als Aufnahmen mit den eigenen Ohren. Dies deckt sich auch mit der Theorie aus Abschnitt 3.4. [MHJS99]

In einer weiteren Versuchsreihe wurde gezeigt, dass es durch sorgfältige Selektion möglich war eine Person zu finden, mit deren Hilfe Aufnahmen mit einer höheren durchschnittlichen Lokalisationsschärfe gemacht werden konnten, als mit einem Kunstkopf. Dies wurde realisiert in dem aus 30 Versuchspersonen die zwei Besten ausgewählt wurden. In Abbildung 15 sind diese beiden Echkopfaufnahmen mit "AVH" bzw. "DOL" gekennzeichnet. Wie zu erkennen ist weisen sie durchschnittlich deutlich weniger Fehler auf, als die Kunstkopfaufnahmen. Besonders auffallend sind die Ergebnisse bei den "within-cone errors", bei denen die Echköpfe überdurchschnittlich gut abschnitten. Zum Zeitpunkt der Studie war dies eine ernüchternde Aussage über die Qualität von Kunstkopfsystemen. [MOCM01]

In Abbildung 15 ist zu erkennen, dass die besten Testergebnisse in allen Grafiken beim "real-life"-Szenario erzielt wurden, also beim direkten Wiedergeben der Testsignale. Die Echkopfaufnahmen (AVH, DOL) schnitten durchschnittlich etwas besser ab, als es



- a) real-life (direktes Testsignal)
- b) Knowles Electronics. Inc. - KEMAR
- c) Georg Neumann GmbH - KU100
- d) Brüel & Kjær A/S - 4100
- e) Head Acoustics GmbH - HMS II
- f) Cortex Electronic GmbH - MK1
- g) Aachen University - ITA
- h) Aalborg University - VALDEMAR
- i) AVH (Echtkopfaufnahme)
- j) DOL (Echtkopfaufnahme)

Abbildung 15 – Auftretende Lokalisationsfehler unterteilt in "out-of-cone error", "within-cone error", "median-plane error", "distance error".

Vergleich von direktem Signal mit sieben Kunstkopfsystemen und zwei Echtkopfaufnahmen. [MOCM04]

die Aufnahmen mit den Kunstkopfsystemen taten. Es ist außerdem zu erkennen, dass die Kunstköpfe unterschiedlicher Bauart auch ihre Vorteile in unterschiedlichen Bereichen haben. So hat der an der Universität Aalborg entwickelte Kunstkopf VALDEMAR beispielsweise die niedrigste Anzahl an "median-plane errors", als auch an "within-cone

errors". In den anderen beiden Kategorien ist dieses Modell aber nur im Mittelfeld. Alles in allem lässt sich erkennen, dass es durchaus Unterschiede in der Lokalisationsfähigkeit mit Aufnahmen der getesteten Kunstkopfmodelle gibt und die Kunstkopftechnologie zum Zeitpunkt der Studie im Bezug auf Lokalisationsschärfe im Durchschnitt nicht mit gut selektierten Echkopfaufnahmen mithalten kann. [MOCM04]

5 Diskussion und Ausblick

Wie in Abschnitt 4.8 berichtet wurde, gibt es bei derzeitigem Stand der Technik noch keine Möglichkeiten die akustische Umgebung mit binauralen Aufnahmesystemen ohne Qualitätsverluste aufzuzeichnen und wiederzugeben. Die Tatsache, dass der Mensch das Hören mit den eigenen Ohren bzw. mit der eigenen HRTF erlernt, macht es extrem schwierig eine allgemeine Lösung zu finden, die es möglichst vielen Personen ermöglicht ein realitätsnahes binaurales Hörerlebnis der aufgenommenen Schallereignisse wahrzunehmen. Eine Möglichkeit zu diesem authentischen, kompromisslosen Hörerlebnis zu kommen, könnte beispielsweise über eine Individualisierung des Aufnahme- und Wiedergabesystems erfolgen, was in der Praxis jedoch nur mit hohem Zeit- bzw. Kostenaufwand umgesetzt werden kann. Bei Kunstkopfsystemen gibt es noch großes Verbesserungspotential, da gezeigt wurde, dass durch Auswahl einer geeigneten Testperson für eine Echtkopfaufnahme bessere Ergebnisse in der Lokalisierung von Testsignalen erzielt werden konnten, als mit den besten verfügbaren Kunstkopfsystemen.

Je nach Anwendungsgebiet kann sowohl die Individualisierung, als auch die Standardisierung von Kunstkopfsystemen einen Weg zur Verbesserung im Bezug auf die Qualität, sowie der Vergleichbarkeit von Aufnahmen liefern. Da in Abschnitt 4.8 gezeigt wurde, dass es große Unterschiede bei den Messergebnissen unter Kunstköpfen gibt, wäre die Standardisierung der geometrischen Maße eines Kunstkopfes ein Weg um den Vergleich von Messungen, verschiedener Kunstkopfsysteme zu verbessern. Vorallem der Ohrenabstand, aber auch der Abstand der Schulter zum Ohr, sowie die Ausformung des Außenohres (Breite, Tiefe und Drehung des Ohres) beeinflussen die Qualität einer Messung am stärksten [FV08].

Alle Anwendungsgebiete, bei denen Standardmessverfahren eingesetzt werden - wie etwa in der industriellen Messtechnik - könnten hier von einer Standardisierung profitieren.

Die Wahl der geometrischen Maße eines Kunstkopfes sind für die Qualität des Systems von ausschlaggebender Bedeutung, jedoch sind sie nicht das einzige Mittel um die Qualität des Aufnahme- und Wiedergabesystems zu verbessern. Eine Verbesserung kann etwa auch durch Methoden im Bereich der Signalverarbeitung erzielt werden. Eine Option binaurale Aufnahmen zu verbessern ist beispielsweise die in Kapitel 4.6 angesprochene Methode der Aufnahme mit Hilfe von Mikrofonarrays, welche in Kombination mit signalverarbeitungstheoretischen Berechnungsschritten zu einer Verbesserung der Ergebnisse führt. Diese Verbesserungen zeigen sich in Form von einer verbesserten Abbildungsschärfe und einer damit verbundenen erhöhten Lokalisationsfähigkeit der hörenden Person.

Durch die stetig steigende Popularität von Kopfhörern, sowie von virtuellen Umgebungen wird die Binauraltechnik in Zukunft eine zentrale Rolle für den Endverbraucher spielen. Auch die zunehmende Rechenleistung von mobilen Geräten ermöglicht der Technologie neue Möglichkeiten der binauralen Wiedergabe. Ein gutes Beispiel ist hierfür das von Apple Music eingeführte Format des Spatial Audios, welches seit 2021 auf mobilen Geräten verfügbar ist. Die bereits angesprochene Individualisierung der HRTF könnte in Zukunft auch hier eine wichtige Rolle spielen. Diese hat ihr größtes Optimierungspotential allerdings vermutlich nicht bei der Verbesserung des Aufnahmeprozesses, sondern in der Aufbereitung von aufgenommenen Signalen mit Hilfe der Signalverarbeitung.

Literatur

- [BB08] J. Blauert and J. Braasch, “Räumliches Hören,” in *Handbuch der Audio-technik*. Springer, 2008, pp. 87–121.
- [Bra02] J. Braasch, “Lokalisationsmodelle des binauralen Hörens,” *FORTSCHRITTE DER AKUSTIK*, vol. 28, pp. 755–760, 2002.
- [DFF⁺07] P. Daniel, H. Fastl, T. Fedtke, K. Genuit, H.-P. Grabsch, T. Niederdränk, A. Schmitz, M. Vorländer, and M. Zollner, “Kunstkopftechnik - eine Bestandsaufnahme,” in *ACUSTProc. Int. Congress on Acoustics ICA/acta acustica/Nuntius Acusticus*, vol. 93, no. 1, 2007, pp. 58–Seiten.
- [FV08] J. Fels and M. Vorländer, “Neue Kunstköpfe,” in *Fortschritte der Akustik - DAGA 2008*, U. Jekosch, Ed., 2008, pp. 655–656.
- [HM05] D. Hammershøi and H. Møller, “Binaural technique, Basic methods for recording, synthesis, and reproduction,” *Communication acoustics*, pp. 223–254, 2005.
- [Lin14] A. Lindau, *Binaural resynthesis of acoustical environments. technology and perceptual evaluation*. epubli GmbH, 2014.
- [MG06] M. Makarski and A. Goertz, “Können binaurale Messungen den Klangunterschied zwischen Studiomonitoren quantitativ erfassen?” *Tonmeistertagung, Leipzig*, vol. 24, 2006.
- [MHJS99] H. Møller, D. Hammershøi, C. B. Johnson, and M. F. Sørensen, “Evaluation of artificial heads in listening tests,” *Journal of the Audio Engineering Society*, vol. 47, no. 3, pp. 83–100, 1999.
- [MOCM01] P. Minnaar, S. K. Olesen, F. Christensen, and H. Møller, “Localization with binaural recordings from artificial and human heads,” *Journal of the Audio Engineering Society*, vol. 49, no. 5, pp. 323–336, 2001.
- [MOCM04] —, “Sound localization with binaural recordings made with artificial heads,” in *Proceedings of ICA2004*, 2004, pp. 3651–3654.
- [Møl92] H. Møller, “Fundamentals of binaural technology,” *Applied acoustics*, vol. 36, no. 3-4, pp. 171–218, 1992.
- [Mol03] C. Moldrzyk, “Binaurale Messungen zur Verifikation einer akustischen Entwurfsmethodik für Architekten,” *FORTSCHRITTE DER AKUSTIK*, vol. 29, pp. 656–657, 2003.
- [Moo12] B. C. Moore, *An introduction to the psychology of hearing*. Brill, 2012.
- [Pau09] S. Paul, “Binaural recording technology: A historical review and possible future developments,” *Acta acustica united with Acustica*, vol. 95, no. 5, pp. 767–788, 2009.
- [PC21] E. Pfanzagl-Cardone, *The Art and Science of Surround and Stereo Recording: Including 3D Audio Techniques*. Springer Nature, 2021.
- [PS90] D. R. Perrott and K. Saberi, “Minimum audible angle thresholds for sources varying in both elevation and azimuth,” *The Journal of the Acoustical Society of America*, vol. 87, no. 4, pp. 1728–1731, 1990.

- [RM21] F. Rumsey and T. McCormick, “Sound and recording : applications and theory.” Routledge, 2021, pp. 486–529.
- [Rum01] F. Rumsey, “Spatial audio.” Focal Press, 2001, pp. 10–51.
- [SBPW16] G. Schlenstedt, F. Brinkmann, S. Pelzer, and S. Weinzierl, *Perzeptive evaluation transauraler binauralsynthese unter berücksichtigung des wiedergaberaums*, 2016.
- [The80] G. Theile, “Über die Lokalisation im überlagerten Schallfeld,” Ph.D. dissertation, Technische Universität Berlin, 1980.
- [TMH20] V. Tkaczyk, M. Mills, and A. Hui, *Testing Hearing: The Making of Modern Aurality*. Oxford University Press, 2020.
- [Vor04] M. Vorländer, “Past, present and future of dummy heads,” *Proc. Acústica, Guimarães, Portugal*, pp. 13–17, 2004.
- [WBG⁺17] E. M. Wenzel, D. R. Begault, M. Godfroy-Cooper, A. Roginska, and P. Geluso, “Immersive sound: The art and science of binaural and multi-channel audio,” pp. 40–62, 2017.
- [Wei08] S. Weinzierl, “Handbuch der Audiotechnik.” Springer Science & Business Media, 2008, pp. 586–590.
- [Xie13] B. Xie, “Head-related transfer function and virtual auditory display.” J. Ross Publishing, 2013, pp. 33–42.