

Research Summary

BinK 

Akustik

Kaan Yalim

Inhaltsverzeichnis

ZIEL DER RECHERCHE UND VORGEHEN	3
GRUNDLAGEN DES RÄUMLICHEN HÖRENS	3
WARUM BINAURALES HÖREN MIT “NUR ZWEI SIGNALEN” FUNKTIONIERT	3
KOORDINATENSYSTEM: WIE RICHTUNGEN BESCHRIEBEN WERDEN (AZIMUT/ELEVATION)	4
ITD & ILD: WARUM LINKS/RECHTS GUT FUNKTIONIERT (HORIZONTAL EBENE)	5
SPEKTRALE HINWEISE: WARUM OBEN/UNTEN UND VORNE/HINTEN SCHWIERIGER IST	5
AUßENOHR & GEHÖRGANG: AKUSTISCHE SCHLÜSSELSTELLEN FÜR DIE KONSTRUKTION	6
OHRMUSCHEL-FORM BEEINFLUSST HRTF DIREKT	6
HRTF, HRIR UND BRIR: WAS GENAU GEMESSEN/GESPEICHERT WIRD	7
MIKROFONPOSITION: TROMMELFELL VS. GEHÖRGANG-EINGANG (PRAKTISCHE KONSEQUENZ)	7
BESTEHENDE KUNSTKOPF-SYSTEME UND MESSANSÄTZE (ALS “BAU-/PRAXIS-REFERENZEN”)	8
KUNSTKOPF ALS “KETTE”: AUFNAHME → WIEDERGABE (KOPFHÖRER)	8
STANDARDISIERTE “AVERAGE ADULT” SYSTEME (HATS).....	8
MESSMETHODIK AUS BESTEHENDEN DATENSÄTZEN	9
KOPF-ÜBER-TORSO-ORIENTIERUNG: WARUM SCHULTERN/TORSO NICHT “EGAL” SIND	9
GRENZEN, QUALITÄT UND ABLEITUNGEN FÜR UNSERE BINKO-TESTS	9
KUNSTKOPF VS. ECHTE “IN-EAR” AUFNAHME (REALISMUSGRENZE)	9
KONKRETE ABLEITUNGEN FÜR UNSERE KONSTRUKTION (WAS WIR GEZIELT TESTEN)	10
KURZFAZIT	10

Ziel der Recherche und Vorgehen

Für meine Diplomarbeit brauche ich ein solides, belegbares Verständnis davon, wie Kopf und Ohr Schall verändern, damit bei binauralen Aufnahmen ein räumlicher Höreindruck entsteht. Das Ziel dieser Recherche ist deshalb, die wichtigsten akustischen Prinzipien (Außenohr, Gehörgang, Kopf/Torso-Einfluss, HRTF/HRIR/BRIR) verständlich zusammenzufassen und daraus Konsequenzen für die Konstruktion und Tests unseres Kunstkopfs abzuleiten.

Ich stütze mich dabei auf vorhandene Fachquellen zu binauraler Technik und Kunstkopf-Systemen (z. B. Hammershøi & Møller, FH St. Pölten-Arbeit, FABIAN-Paper, Messreports und Datenblätter).

Grundlagen des räumlichen Hörens

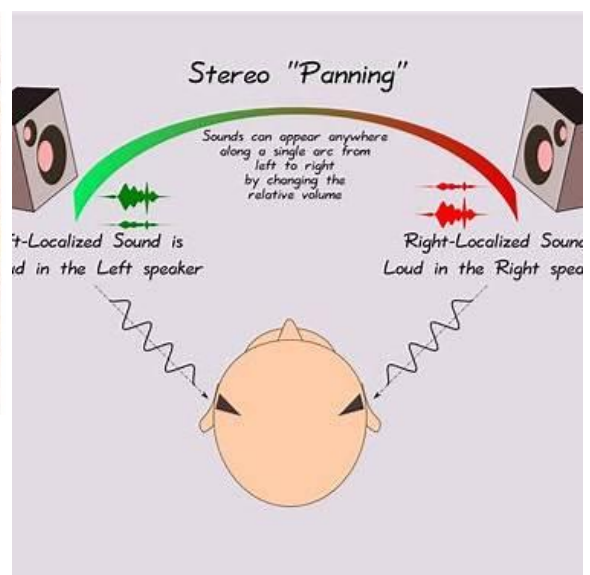
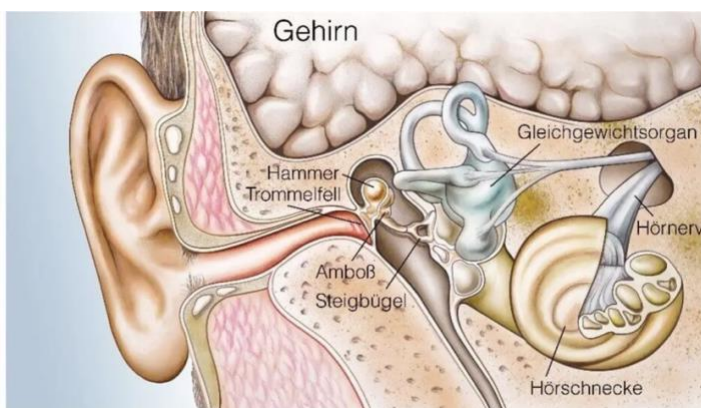
Warum binaurales Hören mit „nur zwei Signalen“ funktioniert

Binaurale Technik baut darauf auf, dass unser räumlicher Höreindruck hauptsächlich aus zwei Eingangssignalen entsteht: dem Schalldruck an unseren zwei Trommelfellen. Wenn diese Signale realistisch aufgenommen und später korrekt wiedergegeben werden, enthält das genug Hinweise („cues“) für räumliches Hören.

Quelle: *2005_Hammershøi_and_Møller.pdf* – Hammershøi & Møller, 2005, PDF-S. 2, Z. 8–9.

Gleichzeitig wird in der Literatur klar gesagt: In der Praxis wird ein echter Mensch beim Aufnehmen oft durch einen Kunstkopf/Manikin ersetzt, der ähnliche akustische Eigenschaften hat. Das ist der Grund, warum man häufig von „artificial-head“ oder „dummy head“ Recording spricht.

Quelle: *2005_Hammershøi_and_Møller.pdf* – Hammershøi & Møller, 2005, PDF-S. 3, Z. 7–8.

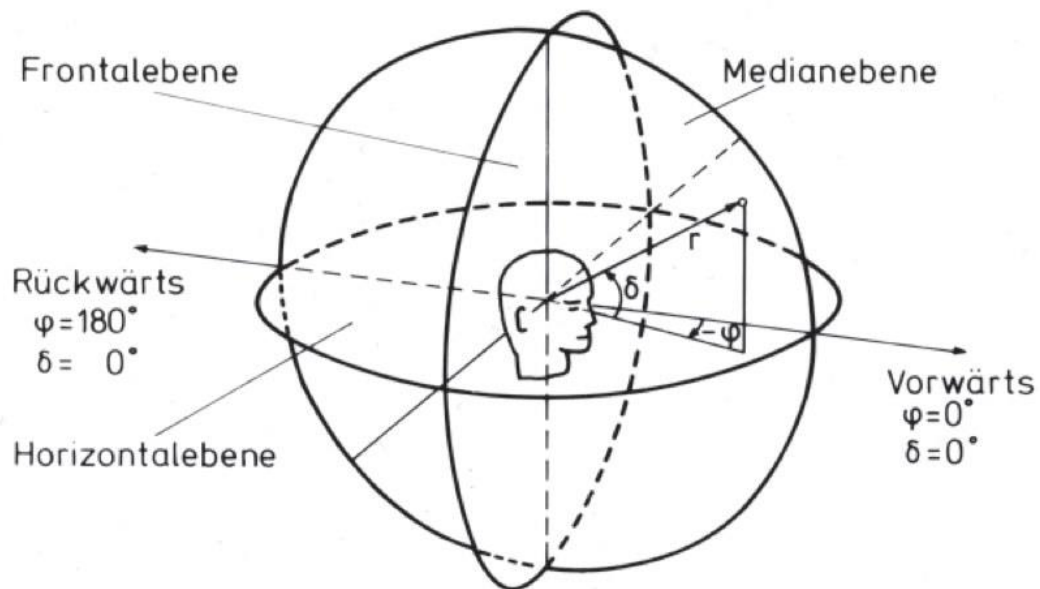


Koordinatensystem: Wie Richtungen beschrieben werden (Azimut/Elevation)

Damit Messungen und Tests vergleichbar sind, wird Schallrichtung oft mit einem Polarkoordinatensystem beschrieben. In der FH-Arbeit wird erklärt, dass der Ursprung auf der interauralen Achse liegt, und es drei Ebenen

gibt: Horizontalebene, Medianebene und Frontalebene. Richtungen auf der Horizontalebene werden als Azimutwinkel φ , auf der Medianebene als Elevationswinkel δ beschrieben.

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 11, Z. 9–20.



ITD & ILD: Warum links/rechts gut funktioniert (Horizontalebene)

Für links/rechts-Ortung (Horizontalebene) wertet das Gehirn Unterschiede zwischen linkem und rechtem Ohr aus. In der FH-Quelle steht explizit, dass dafür vor allem interaurale Pegeldifferenz (ILD) und interaurale Zeitdifferenz (ITD) genutzt werden.

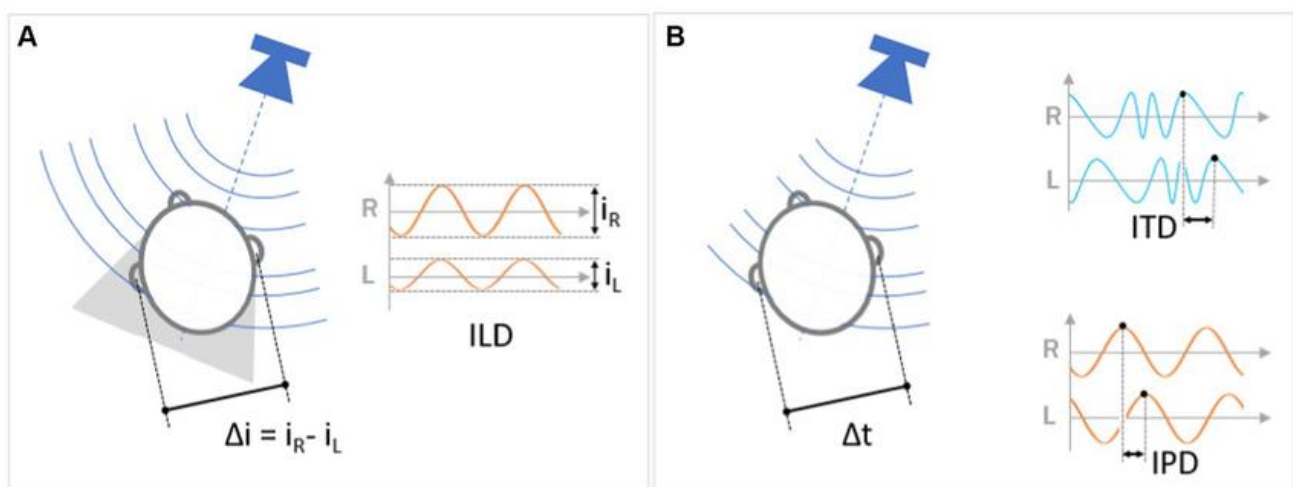
Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 14, Z. 3–6.

ILD (Pegelunterschied) entsteht z. B. durch Kopfabstimmung: Eine Schallquelle links führt dazu, dass rechts weniger Pegel ankommt.

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 14, Z. 7–12.

ITD (Zeitunterschied) entsteht, weil der Schall nicht gleichzeitig an beiden Ohren ankommt (Laufzeitdifferenz Δt).

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 14, Z. 13–14.



Außerdem wird dort beschrieben: Bei Sinussignalen unter ca. 1,6 kHz kann das System Zeitdifferenzen aus der Feinstruktur auswerten; darüber spielen Phasen-/Gruppenlaufzeit-Informationen der Hüllkurve eine größere Rolle.

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 14, Z. 15–18.

Spektrale Hinweise: Warum oben/unten und vorne/hinten schwieriger ist

Für Richtung in der Medianebene (z. B. oben/unten) sind ITD/ILD weniger hilfreich, weil beide Ohren oft ähnlicher angeregt werden. Hier werden vor allem spektrale Veränderungen (Filterwirkung von Ohrmuschel/Gehörgang) wichtig. In der FH-Quelle wird beschrieben, dass bei Breitbandsignalen bestimmte Frequenzanteile je nach Richtung verstärkt/abgeschwächt werden und die dominanten Bänder die Richtung prägen.

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 13, Z. 16–21.

Ein konkretes Beispiel steht dort ebenfalls: Wenn ein Hörereignis viel Anteil um 8 kHz hat, wird es „leicht erhöht“ wahrgenommen (Elevationseffekt).

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 13, Z. 22–24.

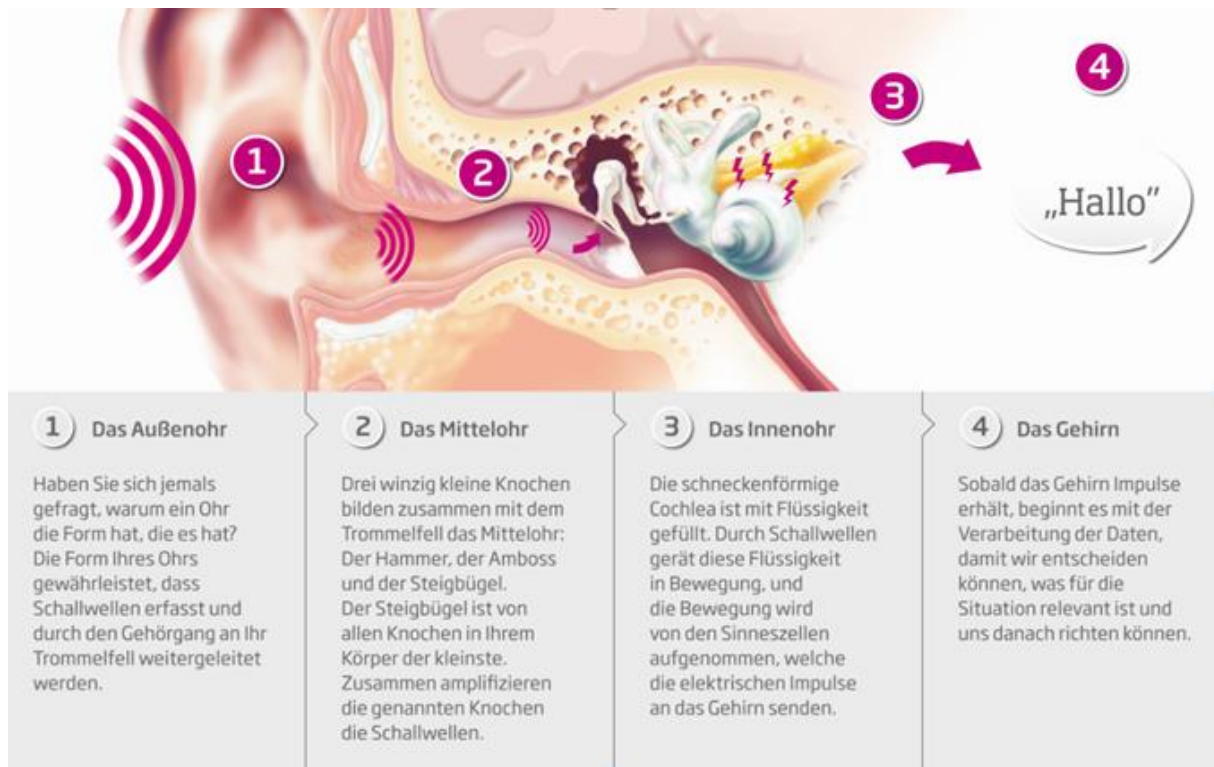
Außenohr & Gehörgang: Akustische Schlüsselstellen für die Konstruktion

Gehörgang-Resonanz und die “4 kHz-Zone”

Die FH-Quelle erklärt: Der Gehörgang (ca. 2–3 cm) beeinflusst den Frequenzgang stark und führt zu einer besonders hohen Empfindlichkeit um ~4 kHz, weil seine Länge ungefähr einem Viertel der Wellenlänge von 4 kHz entspricht (Resonanzprinzip).

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 8, Z. 15–21.

Das ist für den Kunstkopf wichtig, weil der Bereich um 4 kHz im Hören sehr präsent ist. Wenn wir z. B. Gehörgangslänge, Querschnitt oder den “Eingang” stark ändern, kann das die Balance in diesem Bereich verändern.



Ohrmuschel-Form beeinflusst HRTF direkt

In derselben FH-Quelle wird klar gesagt: Die Form des Außenohres beeinflusst die kopfbezogene Übertragungsfunktion (HRTF) und ist wesentlich für die korrekte Wahrnehmung von Kunstkopfaufnahmen.

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 8, Z. 23–24.

Das passt auch sehr gut zu moderneren 3D-Print-Arbeiten: O'Connor & Kennedy untersuchen ausdrücklich, wie Fertigungsstrategie und Pinna-Material die HRTF verändern. Sie betonen außerdem, dass das Ohr die HRTF besonders ab ca. 4 kHz und darüber beeinflusst.

Quelle: *1-s2.0-S0003682X20307143-main.pdf* – O'Connor & Kennedy, 2021, PDF-S. 2, Z. 45–46 und Z. 70–72.

HRTF, HRIR und BRIR: Was genau gemessen/gespeichert wird

Begriffe (kurz und sauber)

HRTF (Head-Related Transfer Function): Frequenzgang/Übertragungsfunktion vom Schallfeld bis zum Ohr (richtungsspezifisch).

HRIR (Head-Related Impulse Response): Impulsantwort, die zur HRTF gehört (Zeitbereich).

BRIR (Binaural Room Impulse Response): Binaurale Impulsantwort inkl. Raum (Reflexionen etc.).

Die DTU-Messarbeit beschreibt z. B. sehr konkret den Schritt von HRTF zu HRIR: Die HRTF wird per inverse FFT in eine HRIR überführt, danach wird eine Zeitfensterung angewendet, um Reflexionen zu reduzieren, und anschließend wird wieder eine FFT gerechnet (und ggf. in 1/12 oder 1/3 Oktavbändern dargestellt).

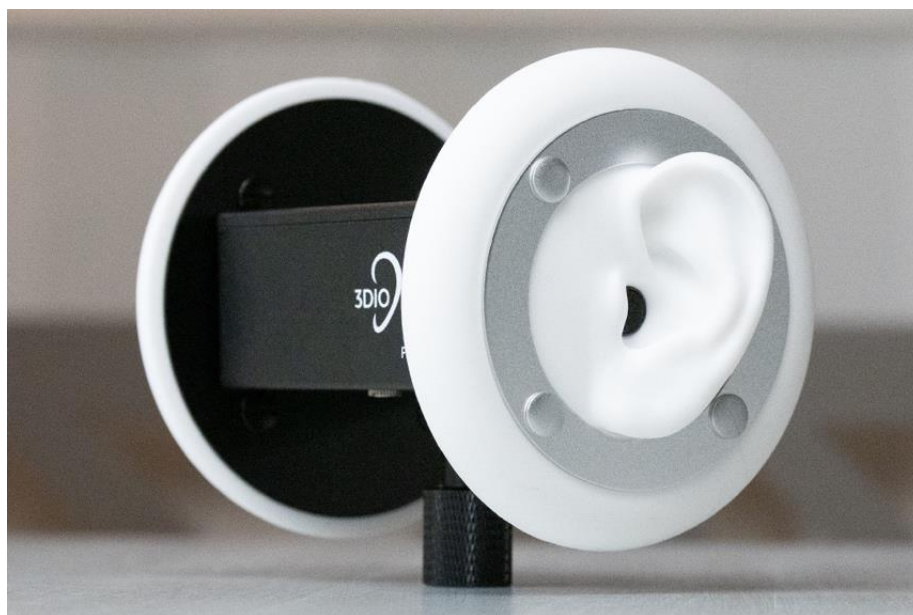
Quelle: *HRTF measurements.pdf* – (DTU Bericht), PDF-S. 10, Z. 16–22.

Mikrofonposition: Trommelfell vs. Gehörgang-Eingang (praktische Konsequenz)

Ein spannender Punkt für den Bau: In vielen Systemen wird nicht direkt am Trommelfell gemessen, sondern am Gehörgang-Eingang/Concha-Bereich. Im FABIAN-Paper steht sehr konkret, dass die Mikrofone am Boden der Cavum Conchae am Beginn des Gehörgang-Eingangs sitzen. Außerdem wird dort behauptet, dass von diesem Punkt bis zum Trommelfell keine Richtungsabhängigkeit der Transferfunktionen existiert (mit Verweis auf weitere Literatur).

Quelle: *Lindau_2006b_FABIAN...pdf* – Lindau & Weinzierl, 2006, PDF-S. 3, Z. 19–20.

Das ist für unser Projekt relevant, weil es begründet, warum Mikrofone am “Ear-Canal-Entrance” (praktischer und robust) häufig eine sinnvolle Wahl sind.



Bestehende Kunstkopf-Systeme und Messansätze (als “Bau-/Praxis-Referenzen”)

Kunstkopf als “Kette”: Aufnahme → Wiedergabe (Kopfhörer)

Hammershøi & Møller betonen, dass binaurale Wiedergabe meist über Kopfhörer geschieht, weil man dort saubere Kanaltrennung und Isolation hat – und dass für “authentische Reproduktion” die gesamte Aufnahme-/Wiedergabekette korrekt kalibriert sein muss.

Quelle: *2005_Hammershøi_and_Møller.pdf* – Hammershøi & Møller, 2005, PDF-S. 3, Z. 9–10.

Sehr praxisnah formuliert es auch Neumann beim KU 100: Der Kunstkopf ähnelt einem menschlichen Kopf, hat Mikrofonkapseln in den Ohren und vermittelt beim Hören über hochwertige Kopfhörer die Illusion, “direkt am Ort” zu sein.

Quelle: *924734.pdf* – Neumann, KU 100 Manual, (Dokumentjahr im Manual), PDF-S. 2, Z. 26–27.

Standardisierte “Average Adult” Systeme (HATS)

Für Mess- und Vergleichbarkeit gibt es standardisierte Head-and-Torso-Simulatoren (HATS). Brüel & Kjær beschreibt seine HATS als Manikin mit kalibrierten Ohrsimulatoren, das eine realistische Reproduktion der akustischen Eigenschaften eines durchschnittlichen erwachsenen Kopf-/Torso-Systems liefert (inkl. Standardbezug).

Quelle: *bp0521.pdf* – Brüel & Kjær, (Data Sheet), PDF-S. 1, Z. 38.

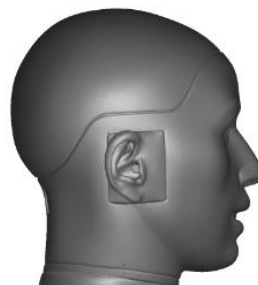
Auch HEAD acoustics beschreibt beim HMS II.3, dass Pinna und Ohrsimulator Anatomie und Performance des menschlichen Außenohrs replizieren und dabei standardkonform sind (IEC/ITU-T P.57).

Quelle: *HMS-II.3-Artificial-Head-1703-Data-Sheet.pdf* – HEAD acoustics, (Data Sheet), PDF-S. 3, Z. 10–14.

Für unser Projekt sind diese Datenblätter hilfreich, um zu argumentieren, dass wir uns beim Maßstab/Grundform am “Average Adult” orientieren können (auch wenn wir selber modellieren), weil es in der Industrie ein verbreiteter Referenzansatz ist.



(a) Whole body



(b) Detailed view of the right ear.

Messmethodik aus bestehenden Datensätzen

Der RWTH-KEMAR-Report beschreibt eine typische HRTF-Messidee sehr klar: Man misst die richtungsspezifische Übertragungsfunktion vom Lautsprecher zu zwei Mikrofonen an den Gehörgang-Eingängen und macht zusätzlich eine Referenzmessung ohne Kopf (am Zentrum), um Einflüsse von Lautsprecher/Mikrofon in der Kette zu berücksichtigen.

Quelle: *Technical_Report.pdf* – (RWTH KEMAR Technical Report), PDF-S. 1, Z. 42–50.

Kopf-über-Torso-Orientierung: Warum Schultern/Torso nicht “egal” sind

Brinkmann et al. zeigen, dass die Kopf-über-Torso-Orientierung (Head-Above-Torso Orientation, HATO) hörbar relevant sein kann. In ihrem Abstract steht z. B., dass bei torso-basiertem Interpolieren eine mediane Hörschwelle der Auflösung von ca. 25° liegt, während head-basiertes Interpolieren ca. 10° als Schwelle zeigt.

Quelle: *brinkmann_etal_2015.pdf* – Brinkmann et al., 2015, PDF-S. 2, Z. 24–29.

Für mich bedeutet das: Wenn man realistische Raum-/Richtungswirkung will, sollte man den Torso/Schultern (oder zumindest eine stabile “Torso-Annäherung”) nicht völlig ignorieren – besonders, wenn es um Dynamik/Bewegung/Orientierung geht.

Grenzen, Qualität und Ableitungen für unsere BinKo-Tests

Kunstkopf vs. echte “in-ear” Aufnahme (Realismusgrenze)

Ein wichtiger Realitätscheck: Minnaar et al. berichten, dass Experimente gezeigt haben, dass Aufnahmen mit Kunstköpfen in der Lokalisation generell schlechter sein können als Aufnahmen, die direkt in den Ohren “gut ausgewählter” Menschen gemacht wurden. Gleichzeitig wird aber auch gesagt, dass manche neuere Kunstköpfe nahe herankommen, und dass es signifikante Unterschiede zwischen verschiedenen Kunstkopf-Modellen gibt.

Quelle: *2004_Minnaar_et_al_ICA_Kyoto.pdf* – Minnaar et al., 2004, PDF-S. 2, Z. 11–22.

Das ist für unsere Arbeit wichtig, weil es begründet, warum wir nicht “den einen perfekten Kopf” erwarten dürfen, sondern durch Iteration/Tests die Qualität Schritt für Schritt verbessern müssen.

(Abbildung: Tabelle/Diagramm “Kunstkopf-Modelle unterscheiden sich” (notfalls als eigene Darstellung mit Literaturzitat); nach diesem Absatz.)

Konkrete Ableitungen für unsere Konstruktion (was wir gezielt testen)

Aus den Quellen leite ich für unsere BinKo-Entwicklung folgende konkrete Punkte ab:

Ohren/Pinna sind kritisch ab ~4 kHz

Ohrmuschel beeinflusst HRTF stark (FH-Quelle).

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 8, Z. 23–24.

Pinna-Einfluss besonders ab 4 kHz (3D-Print-Studie).

Quelle: *1-s2.0-S0003682X20307143-main.pdf* – O'Connor & Kennedy, 2021, PDF-S. 2, Z. 70–72.

Testidee: Ohrvarianten mit gezielten Änderungen (Helix/Antihelix/Concha-Tiefe), dann HRTF-Vergleich und Hörtest.

Gehörgang-Geometrie beeinflusst Resonanzen (4 kHz)

Viertelwellen-Resonanz/Empfindlichkeit um 4 kHz (FH-Quelle).

Quelle: *arbeit-st-polten-binaural.pdf* – (FH St. Pölten), PDF-S. 8, Z. 19–21.

Testidee: Gehörganglänge/Eintrittsöffnung variieren (in kleinen Schritten) und prüfen, ob sich der “4 kHz-Bereich” in Messung/Hören verändert.

Mikrofonposition am Gehörgang-Eingang ist gut begründbar

FABIAN: Mikrofone am Cavum Conchae / Beginn ear canal entrance.

Quelle: *Lindau_2006b_FABIAN...pdf* – Lindau & Weinzierl, 2006, PDF-S. 3, Z. 19–20.

Testidee: Position “Entrance” vs. etwas tiefer/höher testen (sofern mechanisch möglich) und vergleichen.

Torso/Schultern beeinflussen das Ergebnis messbar/hörbar

HATO-Relevanz und Schwellenwerte in Brinkmann et al.

Quelle: *brinkmann_etal_2015.pdf* – Brinkmann et al., 2015, PDF-S. 2, Z. 24–29.

Testidee: Aufnahmen mit/ohne Schulterständer bzw. mit Torso-Anteil vergleichen.

Kurzfazit

Binaurales Hören entsteht, weil Kopf und Ohr den Schall richtungsabhängig filtern und dadurch in den zwei Ohrsignalen genügend räumliche Hinweise stecken (Hammershøi & Møller). Für links/rechts sind ITD/ILD zentral (FH-Quelle), für oben/unten und vorne/hinten vor allem spektrale Hinweise durch Ohrform und Gehörgang. Besonders wichtig ist dabei der Bereich ab ~4 kHz, weil dort die Ohrmuschel die HRTF stark prägt (FH-Quelle; O'Connor & Kennedy). Gleichzeitig zeigen Vergleichsstudien, dass Kunstköpfe nicht automatisch so gut sind wie in-ear Aufnahmen und sich Modelle deutlich unterscheiden (Minnaar et al.). Deshalb ist unser Ansatz – selbst modellieren, Probedrucke und Tests, austauschbare Ohrvarianten – wissenschaftlich gut begründbar und passt genau zu dem, was die bestehende Forschung als sinnvoll nahelegt.

Quellen

<https://verleih.ustp.at/upload/59afec83b00d6.pdf>

https://vbn.aau.dk/ws/portalfiles/portal/227876329/2005_Hammersh_i_and_M_ller.pdf

https://vbn.aau.dk/ws/portalfiles/portal/227976192/2004_Minnaar_et_al_ICA_Kyoto.pdf

[https://www.researchgate.net/publication/263809387_FABIAN - An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom](https://www.researchgate.net/publication/263809387_FABIAN_-_An_instrument_for_software-based_measurement_of_binaural_room_impulse_responses_in_multiple_degrees_of_freedom)

[https://sofacooustics.org/data/database/aachen%20\(high-resolution%20kemar\)/Technical_Report.pdf](https://sofacooustics.org/data/database/aachen%20(high-resolution%20kemar)/Technical_Report.pdf)

<https://d-nb.info/1195040024/34>