

Messung und Untersuchung der Einflüsse der Head-Torso-Winkel auf die Außenohrübertragungsfunktionen

Markus Guldenschuh

9. April 2008, Graz

Mit Hilfe eines Brüel & Kjaer Kunstkopfes wurden für 10 verschiedene Kopf zu Torso Versätze mit je 15 verschiedenen Elevationswinkel (90° bis -41°) und 36 Azimutwinkel logarithmische Sweeps aufgenommen. Mit den über eine Entfaltung gewonnenen HRIRs sollte vor allem der Einfluss der Schulterreflexionen untersucht werden. Die über alle Schulterdrehungen gemittelte HRIR wird als SSH (Surface Spherical Harmonics) dargestellt. Die energetisch relevanten Schulterreflexionen werden zu diesem Grundgerüst additiv hinzugefügt.

1 Einleitung

HRIRs (Head Related Impuls Responses) beschreiben die Übertragungsfunktion des Ohrs im Zeitbereich als Impulsantwort. Sie stellen sozusagen das System dar, das das Signal einer Schallquelle auf dem Weg zu unserem Trommelfell verändert. Die Veränderungen sind in den HRIRs vor allem als Reflexionen an der Pinna des Ohres und an den Schultern ersichtlich. Diese Reflexionen sind entscheidend für das Richtungshören aus der vertikalen Ebene. HRIRs oder HRTFs (Head related Transfer Functions) werden somit verwendet um Spatialisation zum Beispiel über Kopfhörerwiedergabe zu simulieren.

Bisher liegen nur Daten von Übertragungsfunktionen vor, die zwar den Einfluss des Torsos, nicht aber die Verdrehung des Torsos gegenüber dem Kopf berücksichtigen. Die Aufgabe war es deshalb, auch alle HRIRs bei verdrehter Torsostellung zu ermitteln und zu untersuchen, welchen Einfluss die Schulterdrehungen auf die HRIRs haben.

Es hat sich als vorteilhaft erwiesen, den Einfluss der Schulterreflexionen als Differenzsignal von der gesamten HRIR abzuspalten. Dafür wurden die HRIRs für alle Quellrichtungen über die Head-Torso-Winkel gemittelt, um diese gemittelten HRIRs dann von den jeweiligen Ausgangs-HRIRs abzuziehen. Dadurch können wir zwischen einem Grundgerüst an HRIRs, die unabhängig von der Schulterstellung sind, und dem direkten Einfluss der Schulterreflexionen unterscheiden. Das von der Schulterstellung unabhängige Grund-

2 HRIR

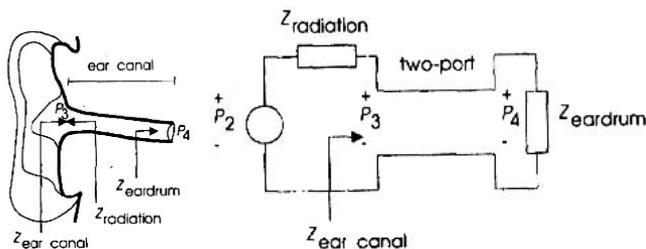


Abbildung 1: Messpunkte im Gehörgang und Ersatzschaltbild (aus Møller)[6]

gerüst wird in SSHs zerlegt. Dieses Modell erlaubt eine einfache Interpolation zwischen den Messpositionen.

2 HRIR

Bei der Außenohr-Übertragungsfunktion können wir zwischen einem richtungsabhängigen Teil und einem richtungsunabhängigen Teil unterscheiden. Vier Messpunkte sind dabei notwendig um den gesamten Gehörgang als System zu erfassen. Es sind dies

- P_1 : Schalldruck im gedachten Mittelpunkt des Kopfes
- P_2 : Schalldruck am Eingang des geschlossenen Gehörgangs
- P_3 : Schalldruck am Eingang des offenen Gehörgangs
- P_4 : Schalldruck am Trommelfell

Der richtungsabhängige Teil ergibt dabei:

$$D(\vartheta, \varphi) = \frac{P_2(\vartheta, \varphi)}{P_1}$$

Der richtungsunabhängige Teil lässt sich als Zweipol am Gehöreingang

$$\frac{P_3}{P_2} = \frac{Z_{earcanal}}{Z_{earcanal} + Z_{radiation}}$$

und als Übertragungsfunktion innerhalb des Gehörgangs $\frac{P_4}{P_3}$ beschreiben. Alle Messpunkte P_2 bis P_4 besitzen die spatialen Informationen, da sie für das linke und für das rechte Ohr separat vorliegen. Es genügt also die Messung von P_1 und eines weiteren Druckes um eine HRIR zu erlangen.[6] In unserem Fall war der zweite Messpunkt der Schalldruck am Trommelfell P_4 , da ein Kunstkopf benutzt wurde, dessen Mikrofonskapseln sich an Stelle des Trommelfells befinden. Als Artefakt wird dadurch jedoch die Resonanz des Gehörganges mit aufgezeichnet.



Abbildung 2: Messaufbau am iem

3 Messaufbau

Im Messlabor des Instituts für Elektronische Musik und Akustik (iem) an der Kunstuniversität Graz befindet sich eine Lautsprecheranordnung mit 15 Lautsprechern, die (annähernd) ein Kreissegment in vertikaler Ausrichtung darstellt.⁽²⁾ Die 15 Lautsprecher reichen, von der Horizontalebene aus gemessen, von $-41,3^\circ$ bis 90° Grad. In der Mitte befindet sich ein Drehteller, der über ein pd-patch angesteuert werden kann.

Es wurden drei Laser, deren Strahl sich im Mittelpunkt der Lautsprecheranordnung kreuzen, am Messplatz montiert. Somit kann die korrekte Positionierung der Messobjekte (in unserem Fall der Kunstkopf) gewährleistet werden. Der Abstand der Lautsprecher vom Messmittelpunkt beträgt im Schnitt 1,5 m.¹

Die Ansteuerung des Drehteller erfolgte von einem PC mit Windows-Betriebssystem aus, während die Messung der Audiodaten selbst über einen Rechner mit Linuxsystem und RME ADI-8 Pro Soundkarten liefen. Um den Messablauf einigermaßen automatisieren zu können² mussten diese zwei Rechner miteinander kommunizieren können. Über die netsend-Befehle in pd war das relativ leicht zu bewerkstelligen. Am Linuxrechner lief ein patch, das die Anregungssignale ausspielte und die Messsignale aufnahm und gleichzeitig korrekt beschriftet abspeicherte. Zudem wurden alle Winkel zu den Laut-

¹Abstandsunterschiede sind zu vernachlässigen, da sie auch in die Referenzmessung mit eingehen und somit aus den Berechnungen fallen.⁽⁴⁾

²Immerhin sollten 5400 Impulsantworten ermittelt werden.

4 Die Messung

Lautspr.	Winkel
1	-41,3°
2	-33,9°
3	-25,3°
4	-17,3°
5	-8,3°
6	0°
7	9,2°
8	18°
9	27,5°
10	37,1°
11	48,1°
12	58,2°
13	69,2°
14	79,6°
15	90,1°

Tabelle 1: Winkel (gemessen von der Horizontalebene)

sprechern gemessen, damit später eine korrekte Beschreibung der HRIRs möglich ist. Dafür verwendeten wir ein Winkelmessgerät mit Laserpointer, das auf einem Stativ im Messmittelpunkt befestigt wurde. Die genaue Anordnung der Lautsprecher kann in der Tabelle (1) nachgelesen werden. Als Messsignal wurde ein logarithmischen Sweep von 2 Sekunden verwendet, der in Matlab erstellt wurde. Die Momentanfrequenz lässt sich folgendermaßen beschreiben[4].

$$\Omega(n) = \omega_{start} \left(\frac{\omega_{end}}{\omega_{start}} \right)^{\frac{n}{N}} \quad (1)$$

Daraus ergibt sich zu jedem Zeitpunkt

$$n(\Omega) = N \frac{\log \left(\frac{\Omega}{\omega_{start}} \right)}{\log \left(\frac{\omega_{end}}{\omega_{start}} \right)} \quad (2)$$

Der Vorteil bei der Messung mit logarithmischen Sweeps ist, dass sie unempfindlich gegenüber Verzerrungen sind, da der Abstand zu den resultierenden Harmonischen stets konstant ist und diese somit leicht eliminiert werden können, und sie zudem eine relativ hohe SNR haben.

4 Die Messung

Für die Messung wurde ein Brüel & Kjaer Kunstkopf 4128 verwendet. Auf einem Klavierhocker positioniert, wurde er mit Hilfe der Laserpointer zentriert. Die Reflexionen

über den Fußboden bzw. über den Klavierhocker können vernachlässigt werden, da diese Reflexionen außerhalb des für uns interessanten Zeitrahmens auftritt und über ein Hanning-Fenster ausgeschnitten werden(5). Die Sensitivity des B.& K. Vorverstärkers wurde auf 12,6 mV/Pa und die Vorverstärkung auf 1 V/Pa gestellt. Die Lautstärke der Sweeps, die auf die Lautsprecher ausspielt wurden, wurden gerade soweit herunter geregelt, dass an den rme-ADCs keine Übersteuerung auftrat. Die Messungen erfolgten bei 44,1 kHz mit 16 bit Auflösung. Insgesamt wurden Messungen für 10 verschiedene Kopf-zu-Torso-Winkel aufgenommen. Diese Winkel reichten von 0° Grad in Zehner-Schritten bis 90° Grad. Pro Kopf-zu-Torso-Winkel ergibt das bei 15 Elevationswinkel³ und 36 Azimuthalwinkel, von 0° Grad in 10°-Grad-Schritten bis 350° Grad, 540 wav-files.

Zum Schluss mussten noch Referenzmessungen am gedachten Kopfmittelpunkt vorgenommen werden. Eine Referenzmessung ist notwendig, um die gemessenen Systemantworten zu entzerren. Wollte man auch den Frequenzgang des Messmikrofons entzerren, müsste die Referenzmessung mit dem selben Mikrofon durchgeführt werden. Da die im Kunstkopf eingebauten Mikrofonkapseln, aber in einem Metallring mit 6 cm Durchmesser eingelassen sind, kommt es durch Druckstau im Bereich von 10 kHz zu einer markanten Erhöhung des Frequenzganges. Deshalb wurde für die Referenzmessung ein GRAS Mikrofon verwendet, das einen ähnlich konstanten Frequenzgang wie die B&K Mikrofone hat. Siehe Abbildung (3). Durch eine Division der Richtungsmessungen durch die Referenz $s = \frac{s(\vartheta, \varphi)}{s_{ref}}$ kürzen sich die Systemeigenschaften, des Messaufbaus, wie etwa die Entfernung der Lautsprecher heraus. Das Referenzmikrofon wurde mit einem Stativ genau am Referenzpunkt der Laser positioniert. Natürlich mussten hier nur mehr die 15 Elevationswinkel aufgenommen werden.

5 Datenauswertung

Nach der Messung lagen als wav-files Sweeps von 2 Sekunden vor. Die Daten sollten in Matlab verarbeitet werden. Sweeps werden entfaltet, in dem ihr Spektrum durch das Spektrum des Originalsweeps (der über die Lautsprecher ausgegeben wurde) dividiert wird. Somit erhält man alle Impulsantworten (dargestellt im Frequenzbereich). Die HRTFs ergeben sich, indem man alle Impulse durch den jeweiligen Impuls der Referenzmessung dividiert.

$$HRTF = \frac{imp(\vartheta, \varphi)}{imp_{ref}(\vartheta)}$$

Diese Division im Frequenzbereich entspricht einer zirkularen Faltung im Zeitbereich. Für Messpunkte, die von der Schallquelle weiter entfernt sind, als der Referenzmesspunkt, bedeutet dies, dass der Anfang des Signals ans Ende verschoben wird. Nachdem man über eine inverse Fourier-Transformation zu den HRIRs gelangt, muss man diese also zuerst zirkular shiften, damit alle Impulse wieder korrekt in ihrem zeitlichen Verlauf dargestellt sind. über ein 128 Samples langes Fenster mit einem 10 Samples langen Hanning-Anfang und einem 60 Samples langen Hanning-Ende wurden die HRIRs auf eine Länge von 2,9 ms reduziert.

³Gegeben durch die Anzahl der zur Verfügung stehenden Lautsprecher

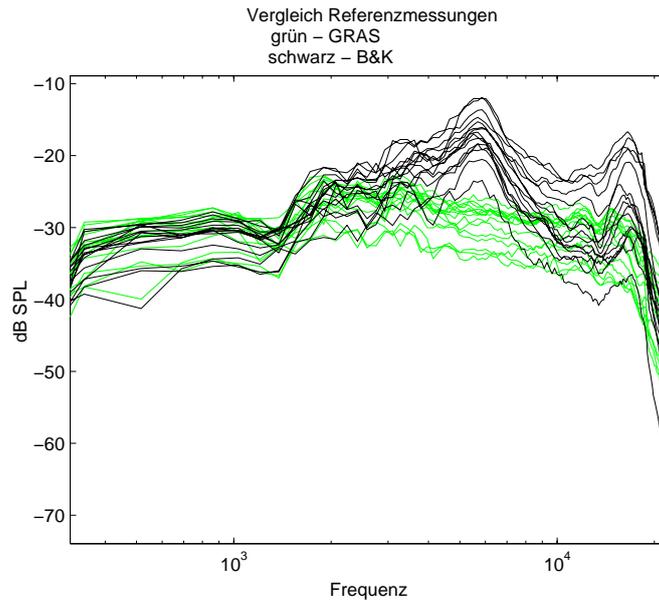


Abbildung 3: Vergleich Referenzmessungen

Es liegen nun also Impulse für jeweils 15 Elevationswinkel ϑ zu je 36 Azimuthwinkel φ bei je 10 verschiedenen Head-Torso-Winkeln δ vor. Grundsätzlich wäre auch eine Darstellung der gewonnenen Daten als HRTFs möglich. Da in dieser Arbeit das Augenmerk aber vor allem auf die Schulterreflexionen gerichtet werden sollte, bot sich die Darstellung im Zeitbereich mehr an, da Reflexionsspitzen im zeitlichen Verlauf leicht erkannt werden.

5.1 Schulterreflexionen als Differenzen dargestellt

Um die Einflüsse der Schulterreflexionen besser bewerten zu können, macht es Sinn, die HRIRs bei verschiedenen Schulterstellungen von einer idealisierten Standard-HRIR in Differenz zu setzen. Diese Standard-HRIR kann als Mittelung aller 10 HRIRs über die Schulterdrehung δ bei gegebenem ϑ und φ verstanden werden. Um diese Mittelung und die anschließende Subtraktion fehlerfrei durchzuführen, müssen die verwendeten Impulse zeitlich exakt aufeinander abzustimmen. Um diese zeitliche Abstimmung möglichst genau zu bewerkstelligen, wurden die Signale mit dem Faktor 18 überabgetastet. Die Standard-HRIR abstrahiert somit den Kopf und einen Hals von der Länge des gesamten Torsos. Diese $HRIR(\vartheta, \varphi)$, die unabhängig von δ ist wird von den $HRIRs(\vartheta, \varphi, \delta)$ abgezogen. Somit ergeben sich Differenzsignale, die schlussendlich wieder additiv zur Standard-HRIR hinzukommen sollen. Siehe dazu auch die Abbildung (4).

5 Datenauswertung

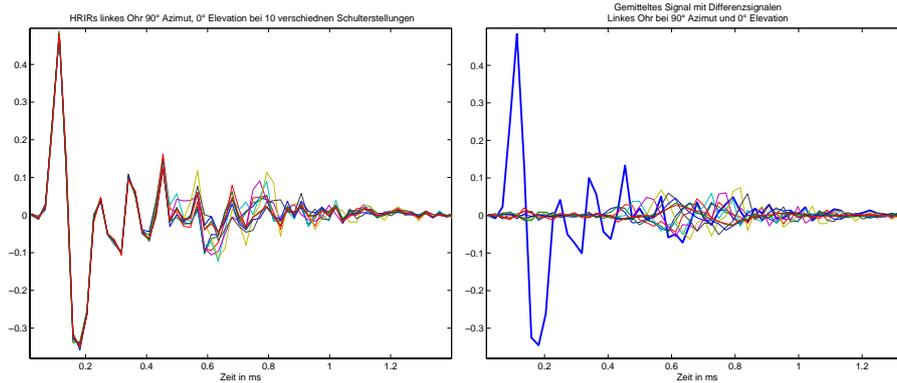


Abbildung 4: Links:HRIR bei 90° Azimut und 0° Elevation. Rechts: Die gemittelte HRIR mit den Differenzsignalen.

5.2 Analyse der Schulterreflexionen

Ein wichtiger erster Schritt ist es, die Leistung der gewonnenen Differenzsignale zu begutachten. Es soll geklärt werden, wie groß die Unterschiede zur gemittelten HRIR sind. Dabei zeigt sich, dass die ersten wichtigen Schulterreflexionen bis zu 14 dB unter der Leistung der gemittelten Signale liegen. Die Einfallsrichtung, die verhältnismäßig starke Reflexionen hervorruft, ist jener Azimutwinkel, der dem betreffenden Ohr am nächsten ist (90° für links, 270° für rechts). Dies ist dadurch zu erklären, dass überhaupt die Signalleistung der HRIRs in diesem Bereich am höchsten ist. Es zeigt sich aber zudem, dass sich der betreffende Azimutwinkel mit der Schulter mitdreht. Bewegt sich die Schulter nach vorne, sind auch die stärksten Reflexionen aus den jeweils weiter vorne liegenden Richtungen zu verzeichnen.

Abbildung (5) zeigt die Bereiche mit den stärksten Differenzsignalen für das linke mit vorwärtsgedrehter Schulter, sowie für das rechte Ohr mit rückwärts gedrehter Schulter. Für die nach vorne gedrehte Schulter kann dieser Bereich durch Elevationswinkel von -20° bis 40° begrenzt werden. Wie schon angemerkt, besteht ein Zusammenhang zwischen Azimut- und Head-Torso-Winkel, der durch folgende Gleichung beschrieben werden kann: $\varphi + \delta = const.$ über diese Achse kann der Bereich eingeschränkt werden.

$$100^\circ \leq \varphi + \delta \leq 140^\circ \quad (3)$$

Bei der nach hinten verdrehten Schulter wird der Bereich in dem die Schallquelle energetisch relevante Differenzen verursacht im Elevationsbereich nur nach oben begrenzt.⁴ $\vartheta < -10^\circ$. Die Head-Torso-Azimut-Achse wird durch folgende Winkel begrenzt:

$$310^\circ \leq \varphi - \delta \leq 350^\circ \quad (4)$$

Allerdings kann in diesem Bereich nicht mehr wirklich von Schulterreflexionen gesprochen werden, da hier die Schultern abschattend wirken und somit Einfluss auf den

⁴Man bemerke jedoch, dass Messdaten nur bis zu -42° vorliegen.

6 Modellierung

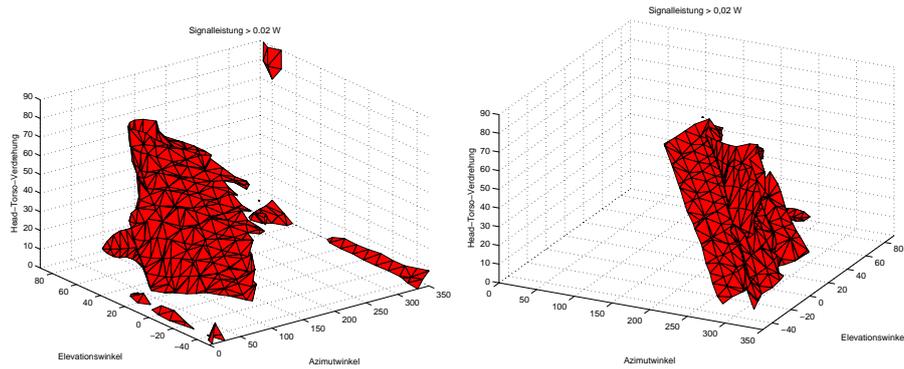


Abbildung 5: Signalleistungen für das linke und das rechte Ohr.

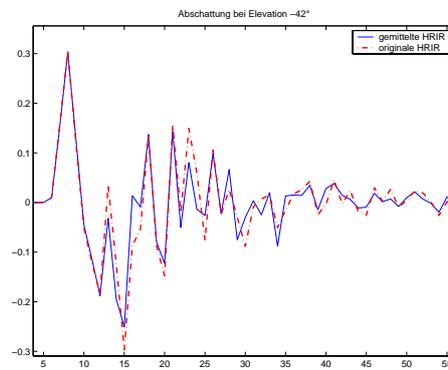


Abbildung 6: Abschattung bei negativen Elevationen und rückwärts gedrehter Schulter

Direktschall haben. Wie man in Abbildung(6) sehen kann betrifft die Abweichung von der gemittelten HRIR den Bereich innerhalb der ersten 15 Samples, die im zeitlichen Bereich der ersten Wellenfront liegen.

6 Modellierung

Es soll ein System erstellt werden, das bei vorgegebenen Eingangswinkeln ϑ und φ die betreffende HRIR erstellt und zudem die Schulterstellung δ in den Bereichen, in denen diese stärkere Reflexionen verursacht, berücksichtigt. Die HRIR wird als SSH dargestellt, was eine Approximation für jeden beliebigen Winkel (oberhalb von -41°) erlaubt. Auch die Modellierung der Schulterreflexionen soll so gestaltet werden, dass zwischen den gemessenen Winkeln interpoliert werden kann. Diese Schulterreflexionen werden dann additiv zum SSH-Grundgerüst hinzugefügt. Siehe dazu auch Abbildung (7)

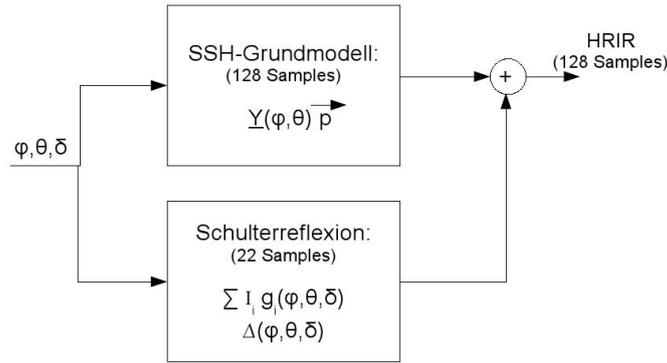


Abbildung 7: Blockschaltbild des analytischen Modells

6.1 Surface Spherical Harmonics

SSHs entsprechen den Lösungen der Laplace Gleichung in Polarkoordinaten. Sie können auch als Fourierzerlegung auf einer Kugeloberfläche verstanden werden. Jede Funktion $f(\vartheta, \varphi)$ kann also als gewichtete Reihe von SSHs beschrieben werden.

$$f(\vartheta, \varphi) = \sum_{nm=1}^{(N+1)^2} c_{nm} Y_{nm}(\vartheta, \varphi) \quad (5)$$

$n=0, \dots, N$ N bezeichnet die Ordnung der SSHs.

$m=-n, \dots, n$

$$nm = n^2 + n + 1 + m$$

Dabei sind c_{nm} die Gewichtungskoeffizienten und Y stellt die SSH-Basiskomponenten dar.

$$Y_{nm}(\vartheta, \varphi) = N_{nm} P_n^m(\cos(\vartheta)) e^{im\varphi} \quad (6)$$

Mit:

$$N_{nm} = \sqrt{\frac{(n - |m|)!(2n + 1)}{4\pi(n + |m|)!}} \quad (7)$$

und

$$P_n^m(z) = -1^m (1 - x^2)^{m/2} \frac{d^m}{dz^m} \times \left[\frac{1}{2^n n!} \frac{d^n}{dz^n} (x^2 - 1)^n \right] \quad (8)$$

Möchte man den gesamten Datensatz also die SSHs zu jedem ϑ und jedem φ zusammenfassen, kann man die Summation in Gleichung (5) auch als Vektormultiplikation anschreiben, wenn man Y als Matrix anlegt. Wenn α die Anzahl der verschiedenen Messpositionen ist, ergibt sich Y zu einer $\alpha \times (N + 1)^2$ Matrix.

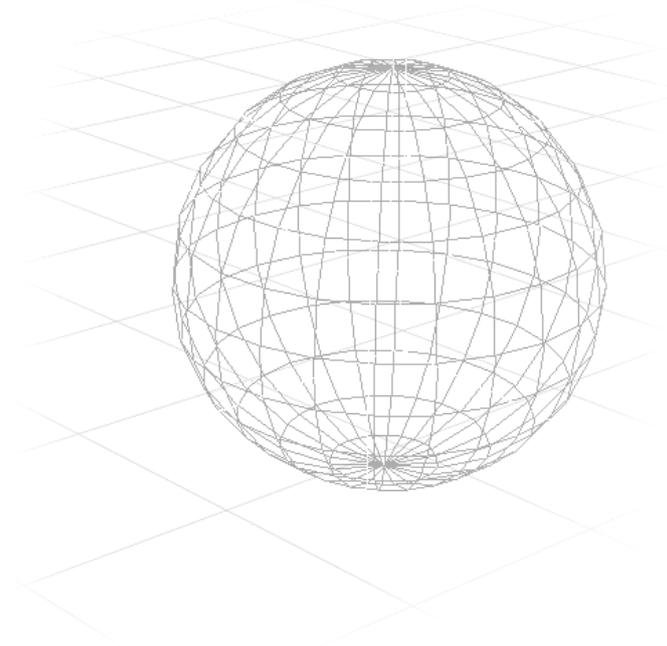


Abbildung 8: Kugel

$$\underline{p}(\vartheta\varphi) = Y(\vartheta\varphi)\underline{c} \quad (9)$$

Für jeden Zeitpunkt t lässt sich so ein Vektor \underline{p} anschreiben, der Informationen zu jeder Messposition beinhaltet. Die Länge des Vektors \underline{c} ist abhängig von der Ordnung des Systems $((N+1)^2)$. Je höher die Ordnung, desto akkurater kann \underline{p} bestimmt werden.

6.1.1 Die Gewichtskoeffizienten

Um die Approximationsmöglichkeiten der SSHs nutzen zu können muss der Koeffizientenvektor \underline{c} bestimmt werden. Dies kann über die LSE-Methode (Least Squared Error) geschehen. Wie bereits beschrieben stellen die SSHs Funktionen auf einer Kugeloberfläche dar. Da unsere Messpunkte kugelsymmetrisch aber nicht äquidistant angebracht wurden, muss der Fehler noch zusätzlich mit den vom Messpunkt abgedeckten Flächenstücken gewichtet werden.

Wie in Abbildung (8) ersichtlich ist, nehmen die Messpunkte um den Äquator (0° Elevation) die größten Flächenstücke ein. Dort muss der Fehler demnach am stärksten bewertet werden. Der Mantel einer Kugelschicht ergibt sich zu

$$M = 2\pi r h \quad (10)$$

Die Höhe h zu

$$h(x) = [\sin(\vartheta(x)) - \sin(\vartheta(x - \Delta))]r \quad (11)$$

6 Modellierung

Der Fehler e lässt sich nun auch als Vektor anschreiben, wenn man die Flächengewichte (die Wurzel aus den Flächenstücken) zu einer Diagonalmatrix W zusammenfasst.

$$\underline{e}_w = W(\underline{p} - Y\underline{c}) \quad (12)$$

Der quadratische Fehler ist demnach:

$$\|\underline{e}\|^2 = (\underline{p} - Y\underline{c})^H W^H W (\underline{p} - Y\underline{c}) \quad (13)$$

Um das Minimum dieser Funktion zu finden wird die Ableitung (nach c) 0 gesetzt. Daraus ergibt sich für c :

$$\underline{c} = (Y^H W^H W Y)^{-1} Y^H W^H W \underline{p} \quad (14)$$

Zur Beschreibung der HRIRs als SSHs sollten möglichst symmetrische Messpunkte vorliegen. Da unsere Messungen nur bis zu einer Tiefe von $-41,3^\circ$ reichen, mussten die restlichen Elevationswinkel interpoliert werden. Dies wurde vorgenommen, indem an die Stelle von -90° das Mittel aller HRIRs bei $-41,3^\circ$ herangezogen wurde. (An den Polen müssen alle Signale zu einem zusammenlaufen.) Die restlichen Winkel dazwischen wurden linear interpoliert.

Je länger der Vector c , also je höher die Ordnung der SSHs (je größer die Matrix Y), desto akkurater die Beschreibung. Bei 15 Elevationswinkeln, läge die maximale Ordnung bei 14.[3] Durch die Erweiterung der Datenpunkte durch die interpolierten HRIRs kann jedoch eine Ordnung von 17 erreicht werden.

6.2 Modellierung der Schulterreflexion

Schulterreflexionen sind, wie in Abschnitt (5.2) beschrieben nur in bestimmten Bereichen von energetischer Relevanz. Für diese Bereiche soll eine additive Korrektur des oben beschriebenen SSH-Grundgerüsts vorgenommen werden. Die Bereiche gliedern sich auf in positive und negative Head-Torso-Verdrehungen. Siehe Abbildung(5). Da sich der Bereich der Quelle, der die energetisch interessanten Schulterreflexionen verursacht mit der Schulter mitdreht, ist es sinnvoll das Modell von einer Achse ($\varphi + \delta = const$) aus anzugehen. Für diese additive Korrektur sind zwei wesentliche Schritte nötig. Die zeitliche Einordnung der Schulterreflexion sowie die Bestimmung der Impulsform selbiger. Die zeitliche Einordnung hängt dabei sowohl vom Azimut-, sowie vom Elevationswinkel und dem Winkel der Head-Torso-Verdrehung ab.

Für die vorwärts gedrehte Schulter variiert der Impulsverlauf im Wesentlichen zwischen den zwei Impulsformen If_1 und If_2 (9). Es werden deshalb diese 22 Samples langen Reflexionsimpulse als Ausgangssignale herangezogen, zwischen denen interpoliert wird. Die Form des Reflexionsimpulses ist abhängig vom Elevations- und Head-Torso-Winkel, sowie von der Head-Torso-Azimut-Achse ($\varphi + \delta = \varsigma$). Die Interpolation wird deshalb über eine lineare Regression 2. Grades bestimmt. Bei der linearen Regression 2. Grades werden sowohl Eingangswinkel, wie auch die quadrierten Eingangswinkel gewichtet, um dann mit einer der Ausgangsimpulsformen multipliziert zu werden.

6 Modellierung

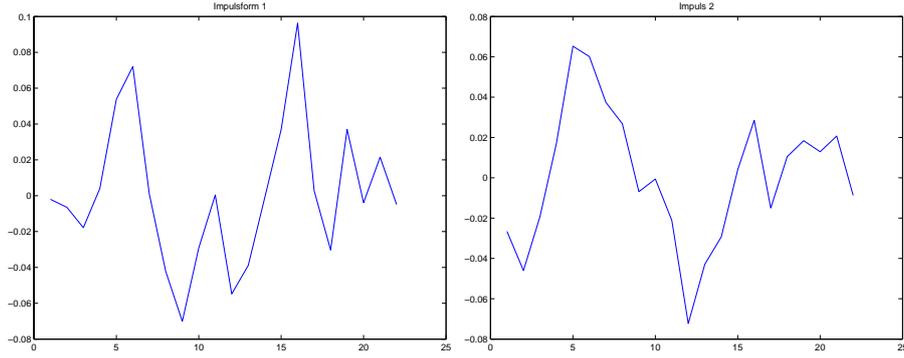


Abbildung 9: Reflexionsimpulsformen If1 und If2 zwischen denen interpoliert wird

$$Imp_{[n]} = (a_1\vartheta + a_2\vartheta^2 + a_3\delta + a_4\delta^2 + a_5\varsigma + a_6\varsigma^2)If_{1[n]} + (b_1\vartheta + b_2\vartheta^2 + b_3\delta + b_4\delta^2 + b_5\varsigma + b_6\varsigma^2)If_{2[n]} \quad (15)$$

Diese Gleichung kann auch in Matrizen-Schreibweise angegeben werden.

$$\begin{pmatrix} Imp_{[1]} \\ Imp_{[2]} \\ \dots \\ Imp_{[22]} \end{pmatrix} = \begin{pmatrix} If_{1[1]} & If_{2[1]} \\ If_{1[2]} & If_{2[2]} \\ \dots & \dots \\ If_{1[22]} & If_{2[22]} \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{pmatrix} \begin{pmatrix} \vartheta \\ \vartheta^2 \\ \delta \\ \delta^2 \\ \varsigma \\ \varsigma^2 \end{pmatrix} \quad (16)$$

Die Koeffizienten a_n und b_n müssen allerdings für alle in Betracht kommenden Winkel ϑ und δ gelten. Die Eingrenzung für den Elevationsbereich wurde im Abschnitt Analyse der Schulterreflexionen(5.2) schon angeführt. Head-Torso-Winkel sind alle von 0° bis 90° in Betracht zu ziehen. Die Summe aus Head-Torso-Winkel und Azimutwinkel ς reicht von 100° bis 140° . Über den LSE-Algorithmus können dann die Koeffizienten a_n und b_n bestimmt werden. Dazu muss folgende Matrix mit allen relevanten Messpunkten $(\vartheta, \delta, \varsigma)$ erstellt werden:

$$A = \begin{pmatrix} -17,3^\circ & -8,3^\circ & \dots & 27,5^\circ & 37,1^\circ \\ (-17,3)^\circ & (-8,3)^\circ & \dots & 27,5^\circ & 37,1^\circ \\ 0^\circ & 0^\circ & \dots & 90^\circ & 90^\circ \\ 0^2 & 0^2 & \dots & 90^2 & 90^2 \\ 100^\circ & 100^\circ & \dots & 140^\circ & 140^\circ \\ 100^2 & 100^2 & \dots & 140^2 & 140^2 \end{pmatrix}$$

Bei 10 Head-Torso-Winkeln und 7 Elevationswinkeln, sowie 5 diskreten Summationswinkel ς ergibt das eine 6×350 große Matrix A.

Um den Fehler e zu bestimmen, müssen auch alle (im vorhin definierten Bereich) gemessenen Reflexionsimpulse in einer Matrix zusammengefasst werden. Dazu müssen diese

6 Modellierung

natürlich alle exakt bei der gleichen Impulsspitze starten. Für diese zeitliche Einordnung wurde zur Steigerung der Genauigkeit mit dem Faktor 8 überabgetastet.

$$\begin{aligned}
 Imp_m &= \begin{pmatrix} Imp_{-17,0,100[1]} & Imp_{-8,0,100[1]} & \dots & Imp_{27,90,140[1]} & Imp_{37,90,140[1]} \\ Imp_{-17,0,100[2]} & Imp_{-8,0,100[2]} & \dots & Imp_{27,90,140[2]} & Imp_{37,90,140[2]} \\ \dots & \dots & \dots & \dots & \dots \\ Imp_{-17,0,100[22]} & Imp_{-8,0,100[22]} & \dots & Imp_{27,90,140[22]} & Imp_{37,90,140[22]} \end{pmatrix} \\
 e &= Imp_m - \begin{pmatrix} If_1[1] & If_2[1] \\ If_1[2] & If_2[2] \\ \dots & \dots \\ If_1[22] & If_2[22] \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{pmatrix} A \quad (17)
 \end{aligned}$$

bzw.

$$e = Imp_m - ICA \quad (18)$$

Das Ergebnis des LSE-Algorithmi lautet:

$$C = (I^H I)^{-1} I^H Imp_m A^H (A A^H)^{-1} \quad (19)$$

Das Ergebnis für C lautet:

$$C = \begin{pmatrix} 0.0063544 & -0.0001770 & -0.0014558 & -0.0000496 & 0.0020511 & 0.0000236 \\ 0.0012377 & -0.0000232 & -0.0007207 & 0.0000471 & -0.0034989 & 0.0000329 \end{pmatrix}$$

Nachdem alle Imp_m zeitlich eingeordnet werden mussten, lag auch für alle diese Schalterreflexionen bereits die Verzögerungszeit in Samples als Matrix D vor. Über das selbe Prinzip der linearen Regression konnten somit auch die Delayzeiten verwertet werden.

$$D = C_{delay} A \quad (20)$$

$$C_{delay} = D A^H (A A^H)^{-1} \quad (21)$$

$$C_{delay} = (0.424403 \quad 0.0005649 \quad 0.013636 \quad 0.00086 \quad 0.1592998 \quad -0.0002745)$$

Für die Modellierung der Reflexionen des rechten Ohres in den Relevanten Bereichen wurde die selbe Methode mit den in (5.2) angegebenen Winkeln angewandt. Da die Differenzsignale in diesem Bereich, aber den Impuls des Dirketschalls betreffen, sind diese Differenzsignale etwas uneinheitlicher als im Bereich der über das linke Ohr aufgenommen wurde. Als Folge dessen musste die Linearkombination mit 4 verschiedenen Ausgangsimpulsformen(10) durchgeführt werden.

Dort lauten die Ergebnisse der Matrizen:

$$C = \begin{pmatrix} -18.6964 & -0.6732 & -4.3086 & 0.0225 & -4.7931 & 0.0175 \\ -59.2077 & -1.0556 & 14.8904 & 0.0902 & -7.1876 & 0.0221 \\ 17.9426 & 0.5085 & -17.2291 & -0.1356 & 7.6700 & -0.0250 \\ 32.0997 & 0.6818 & -0.1762 & -0.0051 & 5.0258 & -0.0117 \end{pmatrix} 10^3$$

$$C_{delay} = (381.6742 \quad 2.0989 \quad 18.0119 \quad 0.4426 \quad 19.7482 \quad 0.1442) 10^3$$

6 Modellierung

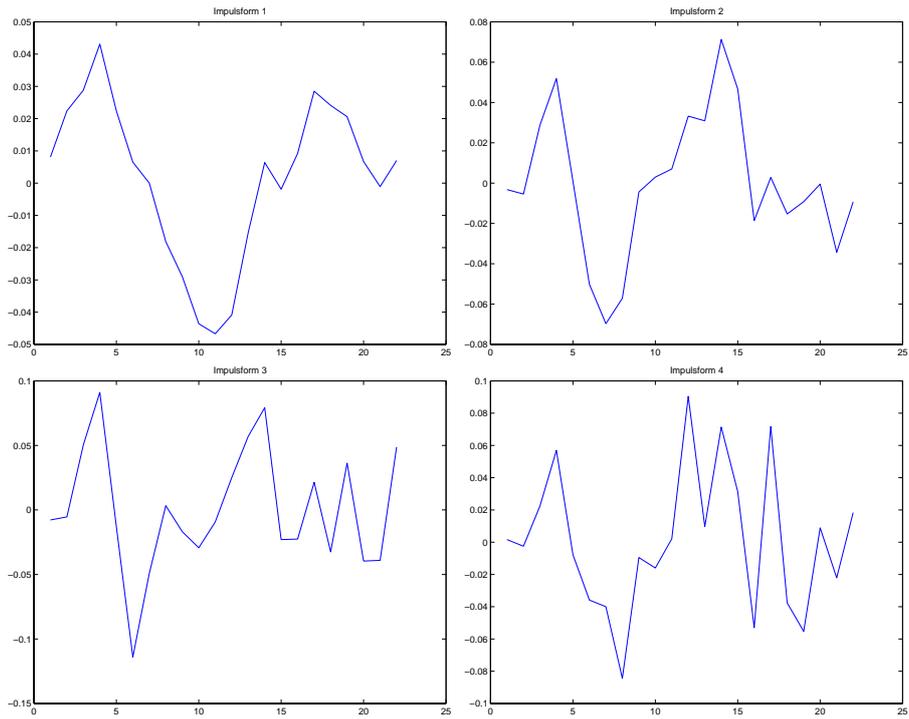


Abbildung 10: Reflexionsimpulsformen zwischen denen interpoliert wird

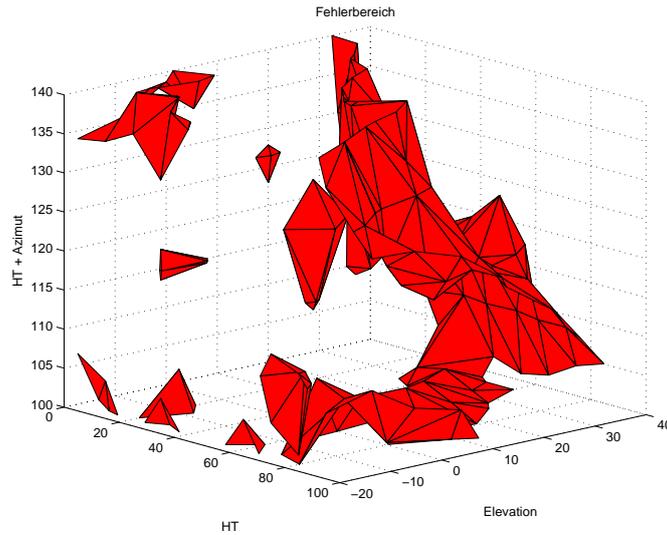


Abbildung 11: Fehlerbereich für die vorwärts gedrehte Schulter

7 Fehleranalyse und Korrekturen

Die Qualität der Modellierung kann evaluiert werden, indem die Abweichung der modellierten HRIR von der gemessenen ermittelt wird. Ein Qualitätsmaß ergibt sich aus dem normierten quadratischen Fehler:

$$Q_2 = 10 \log_{10} \frac{\sum_{n=0}^N |hrir_{orig}[n] - hrir_{model}[n]|^2}{\sum_{n=0}^N |hrir_{orig}[n]|^2} \quad (22)$$

Bei der Modellierung im Bereich von $100^\circ \leq \varphi + \delta \leq 140^\circ$ liegt das Mittel von Q_2 ohne hinzugefügte Schulterreflexionen bei -12,26 dB. Das Modell mit den Schulterreflexionen liegt im Schnitt 0,5 dB darunter. Der Medianwert liegt 0,4 dB darunter. Allerdings gibt es auch Bereiche, in denen das Modell mit hinzugefügter Schulterreflexion vom Original mehr abweicht, als das reine SSH-Gerüst. Dies liegt vor allem an der unzureichend genauen zeitlichen Eingliederung der Schulterreflexionen. Abbildung (11) zeigt den Bereich, in denen das Modell schlecht greift. Durch korrekturen der Delay-time für die Schulterreflexionen im Bereich für hohe Elevationswinkel (und mit zunehmendem δ auch für niedrigere Elevationen) kann Q_2 für das Modell mit hinzugefügten Schulterreflexionen wesentlich verbessert werden. Nun führt das Hinzufügen der Schulterreflexionen in beinahe allen Fällen zu einer Verbesserung des Modells. Das Mittel der Verbesserung liegt nunmehr bei 0,8 dB und der Medianwert liegt 0,6 dB unter dem Medianwert des SSH-Grundgerüsts.

Für die nach hinten verdrehte Schulter, bei denen es für Elevationswinkel unter der

7 Fehleranalyse und Korrekturen

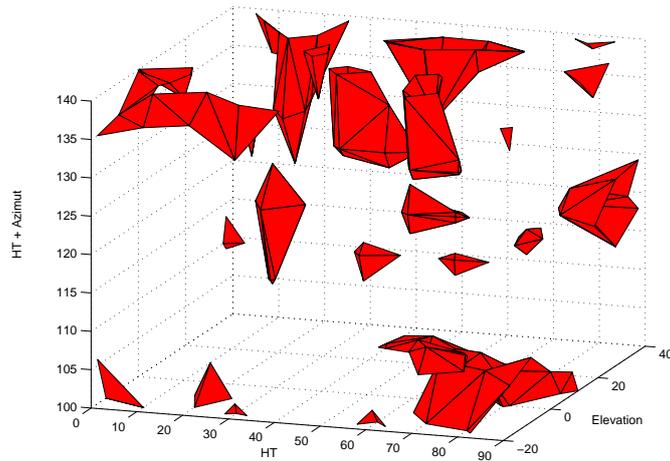


Abbildung 12: verbessertes Q2 für die vorwärts gedrehte Schulter

Horizontalebene zu starken Differenzsignalen gekommen ist, kann durch den Mangel an Auflösung der SSH-Komponenten kein so gutes Ergebnis erzielt werden. Für das SSH-Grundgerüst liegt Q2 bei -8,4 dB. Da unsere Messpunkte keine optimierten räumlichen Abtastpunkte darstellen, liegt die im Endeffekt erreichbare Genauigkeit unter der angenommenen Ordnung 17.[3] Außerdem muss auch räumliches Aliasing berücksichtigt werden. Bei Messabständen von 10° auf einem Kreis, der durch einen Radius von 8 cm bestimmt ist (Abstand vom Ohrmikrofon zum Kopfmittelpunkt), kommen wir zu Kreissegmenten der Länge:

$$\lambda = \frac{2 \cdot 0,08\pi}{36} = 0,014m \quad (23)$$

Demzufolge ist eine fehlerfreie Abtastung nur bis zu einer Frequenz mit der Wellenlänge

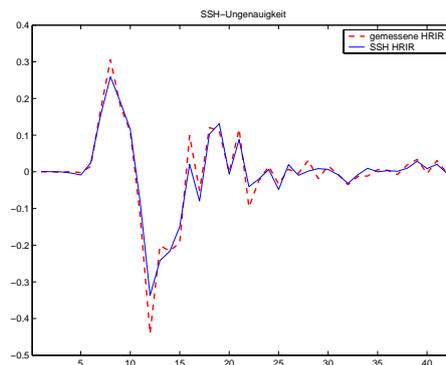


Abbildung 13: Vergleich zwischen der originalen HRIR und der SSH-Approximation

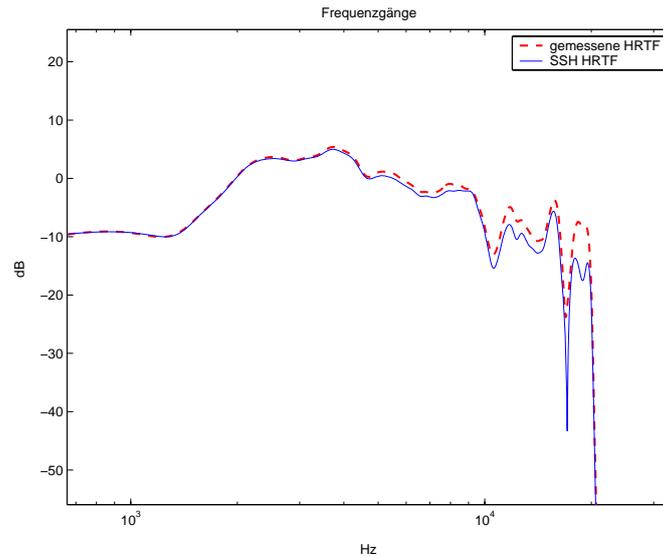


Abbildung 14: Vergleich zwischen der originalen HRTF und der SSH-Approximation

$2\lambda = 0,028m$ gewährleistet. Dies entspricht einer Frequenz von ca. 12 000 Hz. Abbildung (14) zeigt den Vergleich der Frequenzgänge. Oberhalb von 10 000 Hz ist bei der SSH-modellierten HRTF deutlich eine Abweichung erkennbar.

Allerdings führt hier das Hinzufügen der Differenzsignale zu einer stärkeren Verbesserung als im vorigen Fall. Die Verbesserung beträgt im Mittel 1 dB und median 0,8 dB. Dennoch gibt es auch hier Bereiche, in denen die Addition des Korrektursignals zu keiner Verbesserung führt.(15) Auch hier kann das Modell durch eine Korrektur der Delay-Zeit noch weiter verbessert werden, sodass das Modell mit Korrektursignal 1,1 dB mit einem Medianwert von 0,86 dB unter Q2 des Grudgerüstes liegt.(16)

8 Zusammenfassung und Ausblick

Es wurden HRIRs aus 540 Richtungen mit 10 verschiedenen Schulterstellungen aufgenommen. Um den Einfluss der Schultern auf die HRIRs ermitteln zu können, wurden selbige über die verschiedenen Schulterstellungen gemittelt, sodass schlussendlich 540 von der Schulterstellung unabhängige HRIRs vorlagen. Diese gemittelten HRIRs wurden von allen gemessenen HRIRs abgezogen. Dadurch gelangten wir zu Differenzsignalen, die hauptsächlich den Einfluss der Schulterreflexionen widerspiegeln. Es zeigt sich, dass diese Differenzsignale nur in ganz bestimmten Bereichen von energetischer Relevanz sind. Nämlich um 90° und um 270° . Die interessanteste Erkenntnis liegt darin, dass dieser Bereich im Wesentlichen dadurch beschrieben werden kann, dass der betreffende Azimutwinkel der Schallquelle mit dem Winkel der Schulterdrehung bestimmte konstante Werte annimmt. Dies bedeutet, dass, wenn die Schallquelle nach hinten wandert, auch

8 Zusammenfassung und Ausblick

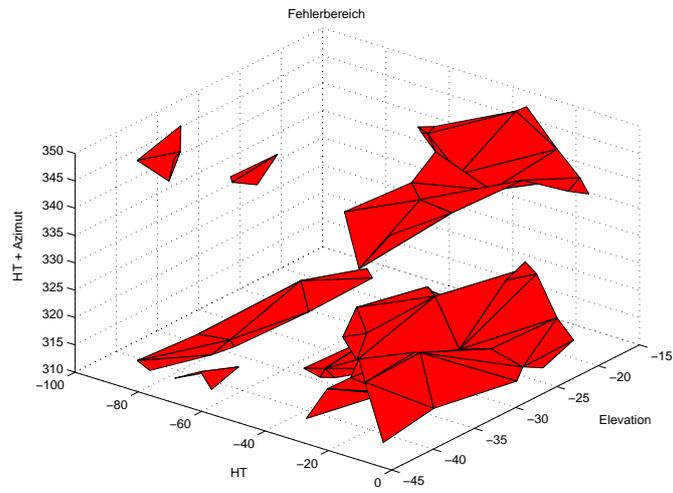


Abbildung 15: Fehlerbereich für die rückwärts gedrehte Schulter

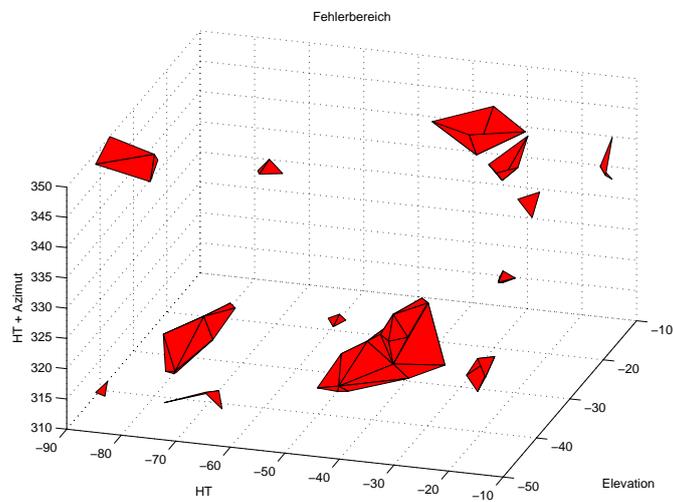


Abbildung 16: Verbessertes Q2 für die rückwärts gedrehte Schulter

die Schulter dementsprechend nach hinten verdreht werden muss, damit sie stärkeren Einfluss auf die HRIR hat. Die gemittelten HRIR werden in weiterer Folge mit Hilfe von SSHs dargestellt, womit sie auch zwischen den Messpunkten approximiert werden können. Die energetisch relevanten Differenzsignale werden mit Hilfe der linearen Regression ebenfalls über Koeffizienten und verschiedene Exemplarimpulse modelliert. Sie können somit zu den SSH-HRIRs hinzuaddiert werden. Dadurch kann die Genauigkeit des Modells im Schnitt um 1 dB erhöht werden. Das System erlaubt eine effiziente Berechnung einer HRIR bei beliebigen Eingangswinkeln ϑ und φ und berücksichtigt zudem die Schulterstellung(7). Im Gegensatz zu einem Look-up-Table, der pro Ohr 540 Einträge für alle 36×15 Quellrichtungen plus 550 weitere HRIRs bei verschiedenen Schulterstellungen benötigte, arbeitet dieses System mit einem SSH Vektor der Länge 324 und 6 Impulsformen (mit einer Länge von nur 22 Samples) für die Modellierung des Schultereinflusses. Damit kann die benötigte Datenmenge auf weniger als ein Drittel reduziert werden.

Um die Wirkung des Systems zu untersuchen, müsste eine Reihe von Hörtest durchgeführt werden. Dann kann entschieden werden, ob das Hinzufügen der Schulterreflexionen tatsächlich auch perceptiv relevant ist und ob die Modellierung genau genug wirkt. Um die Performance des Systems zu erhöhen, müssten jedoch günstiger verteilte Messpunkt vorliegen, um die Ordnung der SSH-Modellierung zu erhöhen.

Literatur

- [1] C. Phillip Brown. Modelling the elevation characteristics of the head-related impulse response. *San Jose State University, Master of Science Thesis*, 1996.
- [2] Corey I. Cheng and Gregory H. Wakefield. Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space. *AES 107th convention*, 1999.
- [3] Michael J. Evans, James A. S. Angus, and Anthony I. Tew. Analyzing head-related transfer function measurements using surface spherical harmonics. *Acoustical society of America*, 1998.
- [4] Angelo Farina. Simultaneous measurements of impulse response and distortion with a swept-sine technique. *AES 108th convention*, 2000.
- [5] Jean-Marc Jot, Veronique Larcher, and Jean-Marie Pernaux. A comparative study of 3-d audioencoding and rendering techniques. *AES 16th international conference*.
- [6] Henrik Møller. Fundamentals of binaural technology. *Appl. Acoust.*, 36:171–218, 1991.
- [7] Henrik Møller, Michael Friis Sørensen, Dorte Hammershøi, and Clemen Boje Jensen. Head-related transfer function of human subjects. *J. Audio Engeneering Society*, 152, No. 5, 1995.