
Messung der eigenen HRTF mit verschiedenen Schulterpositionen

Seminararbeit

Algorithmen in Akustik und Computermusik II, SE

durchgeführt von

Clara Hollomey

Nadine Kroher

Institut für Elektronische Musik
Universität für Musik und darstellende Kunst Graz

Leiter: Franz Zotter, Robert Höldrich

Graz, Januar 2009

Inhaltsverzeichnis

1 Einleitung

2 Die Messung

2.1 Aufgabenstellung

2.2 Der Messaufbau

2.3 Die Messsignalverarbeitung

2.4 Erkenntnisse

3 Weiterführende Versuche

3.1 Projektarbeit Guldenschuh

3.2 Projektarbeit Lindau / Weinzerl

4 Literatur

5 Bilderverzeichnis

Anhang: Die Messergebnisse

1 Einleitung

Die HRTF (head-related transfer function) beschreibt die Übertragungsfunktion des Außenohrs im Frequenzbereich. Im Zeitbereich wird die entsprechende Impulsantwort des Übertragungsweges mit HRIR (head-related impulse response) bezeichnet. Auf dem Weg von der Signalquelle zum Innenohr findet an Torso, Kopf und Pinna sowohl Reflexion als auch Diffraction statt. Abhängig von der Schalleinfallrichtung entstehen unterschiedliche Reflexionsmuster. Durch Überlagerung des Direktschalls mit den oben beschriebenen Reflexionen ergeben sich konstruktive und destruktive Interferenzen. Im Frequenzbereich entspricht dies richtungsabhängigen Frequenzgängen, welche vom Gehirn interpretiert werden. Die HRTF ist also maßgeblich an der Lokalisation von Schallquellen in der menschlichen Gehörwahrnehmung beteiligt. Dieser Effekt tritt in der vertikalen Ebene am deutlichsten auf.

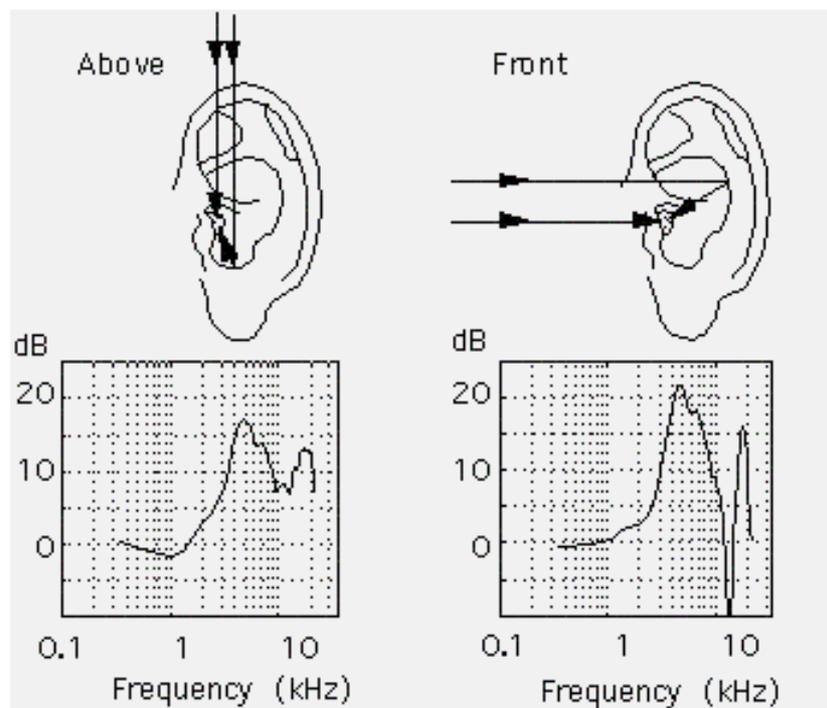


Abbildung 1: Frequenzgang durch Reflexion an der Pinna

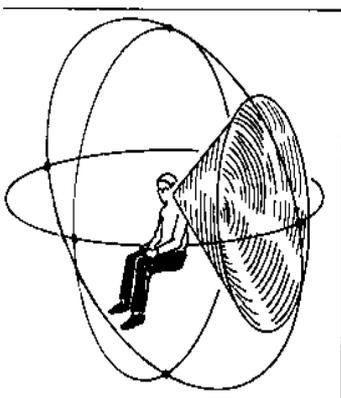


Abbildung 2: Cone of Confusion

Grundsätzlich lässt sich die HRTF in einen richtungsabhängigen und einen richtungs-unabhängigen Teil spalten. Der richtungsabhängige Teil beschreibt die oben genannten Reflexionen an Torso, Kopf und Pinna. Der richtungsunabhängige Teil hingegen beschreibt die Resonanz des Ohrkanals. Aufgrund der von uns verwendeten Messmethode mit Bügel-mikrofonen, welche

im Weiteren noch näher beschrieben wird, war es uns möglich größtenteils nur den richtungsabhängigen Teil zu bestimmen, welcher für die Lokalisationseigenschaften ausschlaggebend ist. Allerdings war der Gehörgang durch die Bügelmikrofone nicht vollkommen verschlossen, wodurch ein Teil der Ohrtunnelresonanz als systematischer Messfehler im Messergebnis enthalten ist.

Es gibt grundsätzlich mehrere standardisierte Verfahren, die nicht-individuelle HRTF zu ermitteln:

- ohne Kunstkopf: Jecklin-Scheibe OSS, Kugelflächenmikrofon
- mit Kunstkopf: Kemar-Messungen als Referenz



Abbildung 3:
KEMAR Kunstkopf

In der HRTF findet sich zusätzlich zu den oben beschriebenen richtungsabhängigen Frequenzgängen, welche durch Interferenzen erzeugt werden auch die Lokalisation von Schallquellen auch anhand von Laufzeit (ITD)- und Intensitätsdifferenzen (ILD) zwischen den beiden Ohren wieder: Der Pegel bei hohen Frequenzen ist am von der Quelle abgewandten Ohr deutlich niedriger, was auf die Abschattung durch den Kopf erklärbar ist (ILD). In der HRIR ist erkennbar, dass das Signal an den Ohren mit unterschiedlicher Verzögerung eintrifft (ITD). An einigen Stellen, die als „cone of confusion“ bezeichnet werden, sind ITD und ILD jedoch identisch. Diese Mehrdeutigkeit kann durch die durch Reflexionen erzeugten Frequenzmuster entscheidend verringert werden.

Aufgrund der unterschiedlichen anatomischen Gegebenheiten sind die HRTFs bei jedem Menschen unterschiedlich. Es entstehen individuelle Frequenzmuster. Versuche zur virtuellen Simulation mit einer „fremden“ bzw. gemittelten HRTF führen zu einer schlechteren Lokalisationsfähigkeit und als mit der individuellen HRTF [1]. Beim „Hören mit fremden Ohren“ tritt allerdings nach längerer Zeit ein Lerneffekt ein und die Lokalisationsfähigkeit verbessert sich [2]. Grundsätzlich beruht die Interpretation der erzeugten Frequenzmustern auf Erfahrungswerten und ist nicht angeboren.

Grundsätzlich kann eine Binauralaufnahme, also eine mittels z.B. Kunstkopf aufgenommene Tonaufnahme nur unter Verwendung von Kopfhörern optimal reproduziert werden, über

Lautsprecher entstünde ein oft hohl klingender Stereo-Effekt. Dies ist dadurch zu erklären dass das Signal bei der Aufnahme durch die HRTF des Kunstkopfs entsprechend gefärbt wird. Bei der Lautsprecherwiedergabe findet diese Färbung durch die HRTF des Hörers, also auf dem Weg vom Lautsprecher zum Innenohr, erneut statt. Die so entstehenden verstärkten Klangfärbungen können nicht mehr korrekt interpretiert werden.

Trotzdem werden seit den 70er Jahren des vorigen Jahrhunderts HRTFs (vorwiegend über Kunstkopf-Aufnahmen) vor Allem bei Hörspielproduktionen als Stilmittel eingesetzt. Vor Allem der Standort des Kunstkopfs kann in die Handlung des Stücks eingebaut werden und erwünschte Stimmungen unterstreichen [3].

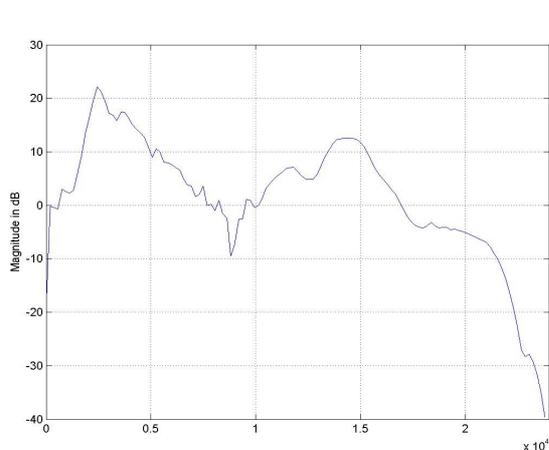
Eine andere praktische Anwendung von HRTFs ist die virtuelle Positionierung von Schallquellen bei der Kopfhörerwiedergabe. Hierbei wird das Signal mit der HRIR der entsprechenden Schalleinfallrichtung gefaltet und somit eine Lokalisation aus dieser Richtung erzeugt. Aufgrund der zuvor beschriebenen Individualität der HRTFs sind dem jedoch Grenzen gesetzt.

3 Die Messung

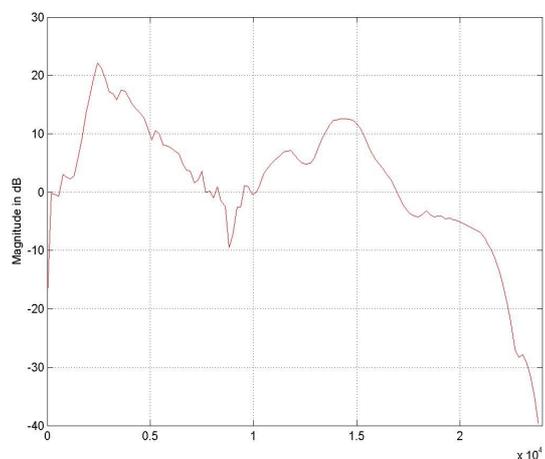
3.1 Aufgabenstellung

Mit dem Versuchsstand im 1. Stock sollen die HRTF einer Versuchsperson evt. mit 3-5 Rotationswinkeln für die Schulterpartie gemessen werden. In einem kleinen Selbstversuch wurde anschließend weißes Rauschen mit den HRIRs gefaltet und die Lokalisationsschärfe bewertet. Als Richtwerte für die gemessenen HRTFs dienten dabei die KEMAR-HRTFs, 1994 durch einen Kunstkopf, der einen durchschnittlichen Menschen repräsentieren soll, ermittelte Übertragungsfunktionen.

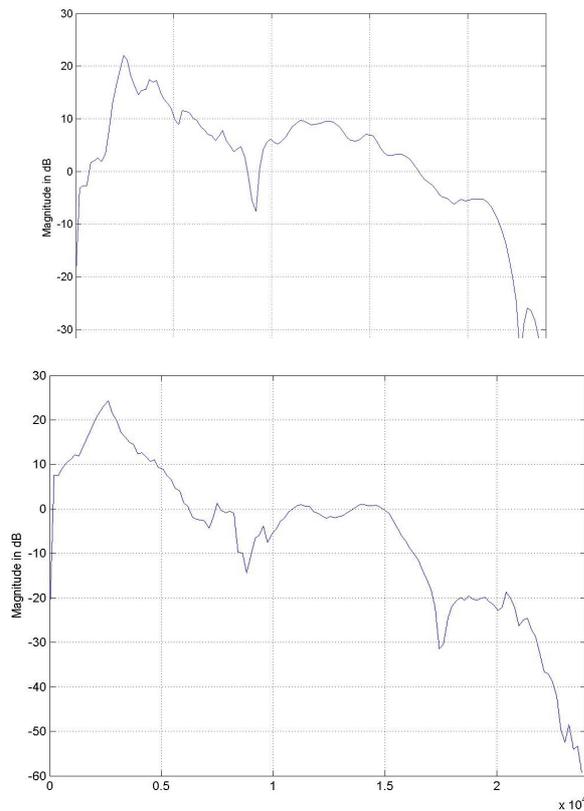
Einige KEMAR-HRTFs:



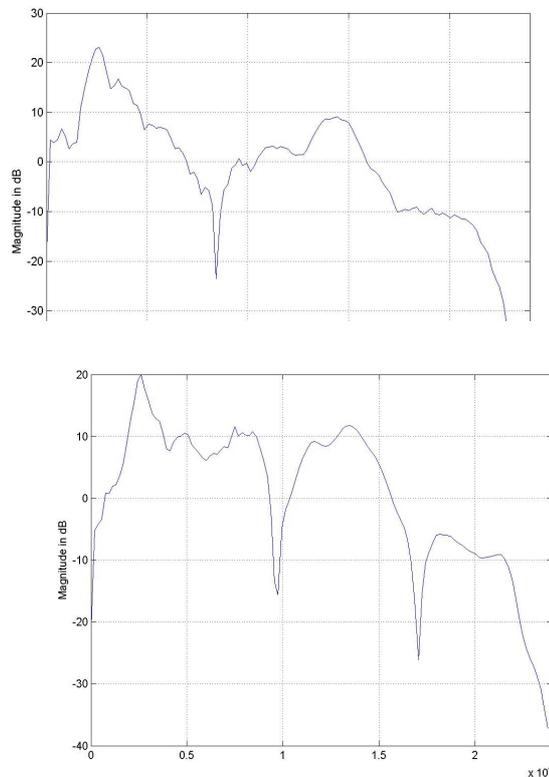
0 Grad links



0 Grad rechts



80 Grad links



80 Grad rechts

3. Der Messaufbau

Zur Messung des richtungsabhängigen Teils muss die Resonanz des Ohrkanals nicht gemessen werden. Wir betrachten also den Weg des Schalls von der Quelle zum Eingang des Gehörgangs. Die Messung erfolgte mittels Kopfbügelmikrofonen von Brüel & Kjaer. Der Aufbau im Messlabor des IEM befindet sich in einem reflexionsarmen Raum um zu verhindern dass stark Raumreflexionen die Messung beeinträchtigen. Im radialen Abstand von ca. 1,5 m vom Messobjekt befinden sich auf einem Kreisabschnitt 15 Lautsprecher im Winkelbereich von -41° bis $+90^\circ$ bezüglich der Horizontalebene. Die Versuchsperson sitzt auf einem Klavierhocker, welche von einem Drehteller rotiert werden kann. Um den Mittelpunkt des Messaufbaus bestimmen zu können, wurden seitlich angebrachte Laser verwendet. Die Wiedergabe der Testsignale, deren Aufnahme und die Ansteuerung des Drehtellers erfolgt über ein PD-patch, welches die aufgezeichneten Messdaten auch

teilweise weiterverarbeitet und benannt. Insgesamt wurden 36 Stellungen am Drehteller eingestellt und es lagen somit insgesamt 540 Ergebnisse pro Messung vor.

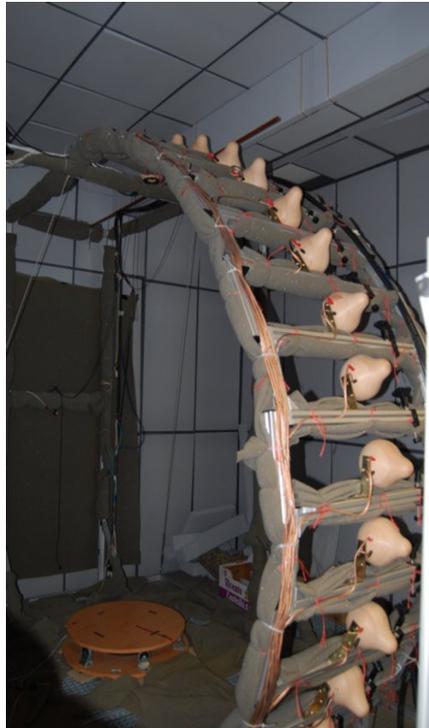


Abbildung 4: Messaufbau am IEM

3.2 Die Messsignalverarbeitung

Die Messung der HRTF erfolgt über die MESM (multiple exponential sweep method, [4]), wobei sich zeitlich überschneidende exponentielle Sinus-Sweeps mit einer Dauer von 2s abgespielt werden. Diese Methode bietet einen verbesserten Signal-Rauschabstand als die direkte Messung mittels Impuls, da mehr bei geringerer Amplitude Energie in den Messaufbau gebracht wird. Durch Multiplikation mit dem inversen Testsignal ergibt sich die Impulsantwort. Hier sind nichtlineare Verzerrungen größtenteils als zeitlich frühere Impulsantworten mit niedrigerer Amplitude erkennbar. Diese können daher im Vergleich zu anderen Methoden mit entsprechender Fensterung eliminiert werden. Eine Fouriertransformation liefert den entsprechenden Frequenzgang. Solange sich die Impulsantworten zeitlich nicht mit den ersten Verzerrungen der vorherigen Impulsantwort überschneiden, können die Testsignale zeitlich überlappend abgespielt werden. Dies bedeutet eine kürzere Messzeit und somit weniger unerwünschte, aber nicht vermeidbare Kopfbewegungen der Versuchsperson während der Messung. Da eventuelle Raumreflexionen deutlich später eintreten als die Reflexionen an Torso, Pinna und Kopf, können diese ebenfalls mit einer entsprechenden Fensterung eliminiert werden.

Um die Übertragungsfunktion des verwendeten Lautsprechersystems heraus rechnen zu können, wurde eine Referenzmessung mit einem Messmikrofon von Brüel & Kjaer durchgeführt und die gemessenen HRTFs später durch diese Übertragungsfunktion dividiert.

Sämtliche Berechnungen wurden in MATLAB durchgeführt, die entsprechenden Plots finden sich im Anhang.

3.3 Erkenntnisse

Die erhaltenen Plots ähneln sehr den als Referenz herangezogenen KEMAR-HRTFS. Nach dem Abspielen der HRIRs als wav.-files ist klar, dass die Lokalisation (auch mit fremden HRTFs) auf jeden Fall funktioniert. Die einzigen Probleme entstehen, wenn die Schallquelle hinter der abhörenden Person positioniert wird. Hier kommt es zu einer Vorne-Hinten-Verwechslung wahrscheinlich wird, bei der Beispielsweise 60° als 30° gedeutet werden können.

Interessante Resultate ergab auch der Versuch, die Abschattung durch den menschlichen Kopf und daraus resultierende Pegeldifferenzen mittels einseitigem Anheben der Lautstärke zu kompensieren. Die Lokalisationsfähigkeit scheint diesbezüglich kaum Abhängigkeiten aufzuweisen.

4 Literaturstudie

Ergänzend zu den von uns durchgeführten Messungen wurden die Projektarbeit „Messung und Untersuchung der Einflüsse der Head-Torso-Winkel auf die Außenohrübertragungsfunktionen“ von Markus Guldenschuh [5], der Konferenzbeitrag „Zur Wahrnehmbarkeit von Diskretisierungen in der dynamischen Binauralsynthese“ von Alexander Lindau und Stefan Weinzierl [6] sowie die Diplomarbeit „Head movements, an additional cue in localization“ von Philip Mackensen an der Technische Universität Berlin herangezogen.

4.1 Prokejtarbeit Guldenschuh [5]

Die Projektarbeit von Markus Guldenschuh wurde mit dem selben Messaufbau im Messlabor des IEM durchgeführt. Anstelle einer realen Versuchsperson verwendete er jedoch einen Kunstkopf von Brüel & Kjaer. Da sich beim Kunstkopf die Mikrofone an der Stelle des Trommelfells befinden, wurde die Resonanz des Ohrkanals mit gemessen. Es zeigte sich, dass die Schulterreflexionen bei verdrehtem Head-Torso-Winkel bis zu 14 dB von einem gemittelten Wert abweichen. Prinzipiell sind die Reflexionen der Schulter bei normaler Schulterposition bei einem Einfallswinkel von 90° bzw. 270° am stärksten. Wenn die Schulter nach von bzw. nach hinten bewegt wird, verändert sich dementsprechend auch die Einfallsrichtung bei der die Schulterreflexionen am stärksten ist. Die Änderung der Schulterstellung hat also einen deutlichen Einfluss auf die Gestalt der HRTF.

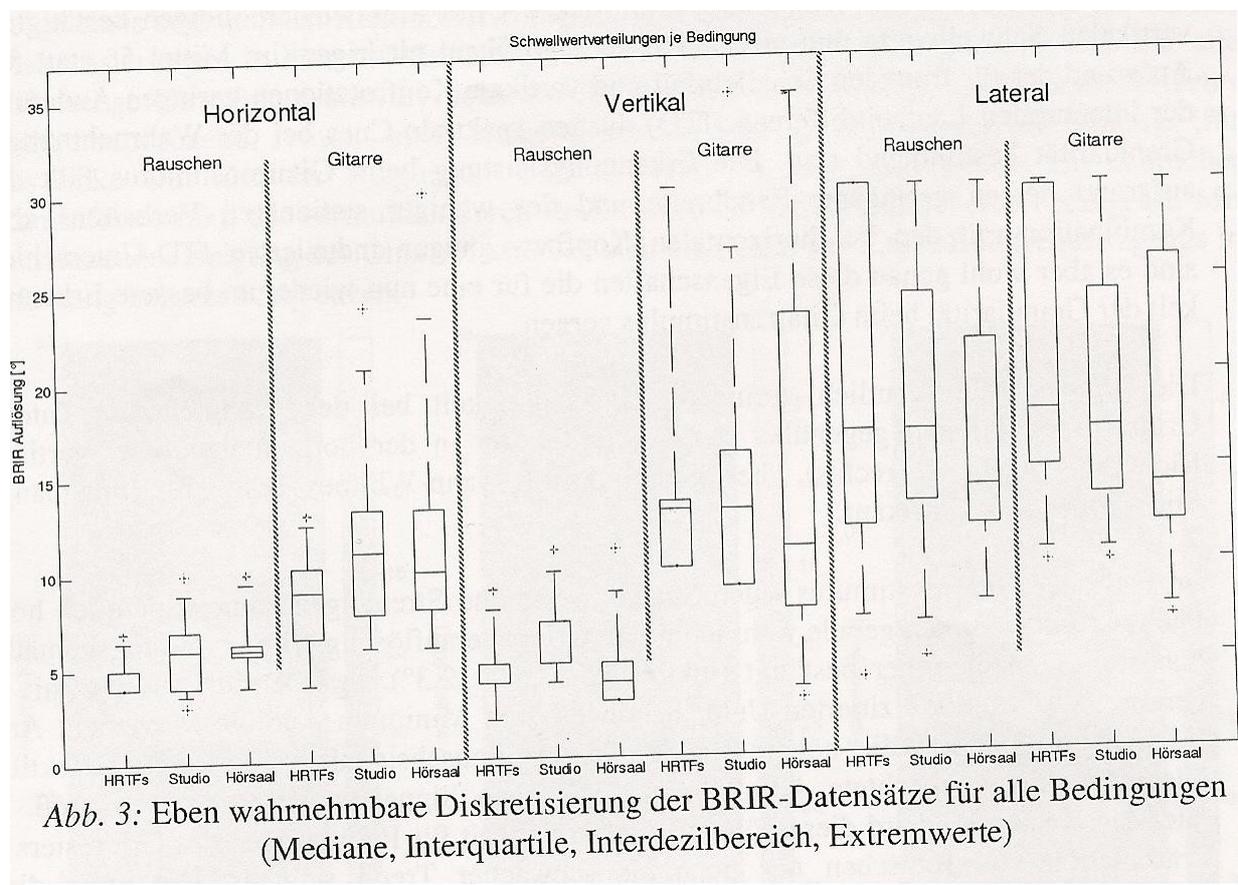
4.2 Konferenzbeitrag Lindau / Weinzerl [6]

Die Arbeit von Alexander Lindau und Stefan Weinzerl wurde an der TU Berlin durchgeführt. Zunächst wurden mit einem möglichst realistischen Kunstkopf- und Torsosimulator die HRTFs für verschiedene Schulterpositionen in verschiedenen Räumlichkeiten (reflexionsarmer Raum, Tonstudio und Hörsaal) aufgenommen. Anschließend wurden zwei Stimuli (rosa Rauschen, Stück für akustische Gitarre) vorgespielt. Hierbei wurden über einen Headtracker die Kopfposition der Versuchsperson ermittelt und so die Impulsantwort an die Kopfposition in Echtzeit angepasst. Es wurde zwischen vertikalen, horizontalen und lateralen Kopfbewegungen unterschieden.

Die grundlegende Frage hierbei war, wie niedrig dabei die Auflösung des Winkelrasters für die Kopfbewegung sein darf, bevor die Versuchsperson einen Unterschied wahrnehmen kann. So wurde die Rasterauflösung schrittweise verkleinert und die dazugehörige Lokalisationsschärfe ermittelt. Als Versuchsumgebungen wurden dabei ein reflexionsarmer Raum, ein Hörsaal und ein Tonstudio verwendet.

Ergebnisse: Für 50% der Versuchspersonen war eine Gridraasterauflösung bei Rauschen von 6° horizontal, 5° vertikal und 16° lateral gerade noch hörbar. Bei dem Ausschnitt aus dem Gitarrenstück waren die Werte für die Rasterauflösung etwas geringer (9° horizontal / 12° vertikal / 16° lateral). Keine Versuchsperson konnte eine Auflösung $< 3^\circ$ detektieren.

Abbildung 5: Graphische Darstellung der Ergebnisse



4.3 Dissertation Mackensen [7]

In dieser Arbeit wird der Einfluß der Kopfbewegung auf die Lokalisationsfähigkeiten untersucht. Das Projekt besteht aus 3 Versuchen:

1. Versuch: Der Kunstkopf ist in einem horizontalem LS-Array mit 27 Phantomschallquellen fixiert. Die Versuchsperson im Nebenraum bekommt ein Signal vom Kunstkopf über Kopfhörer eingespielt und soll die Lokalisation grafisch festhalten.
2. Versuch: Die Versuchsperson darf den Kopf horizontal bewegen. Über einen Head-Tracker wird die genaue Position ermittelt und der Kunstkopf nachgeführt.
3. Versuch: Jetzt befindet sich die Versuchsperson im Raum mit dem LS-Array, die Lautsprecher sind hinter einem Vorhang versteckt und die virtuelle Schallquelle soll lokalisiert werden.

Ergebnisse: Gab es beim ersten Versuch mit statischem Kunstkopf noch Vorne / Hinten-Verwechslungen im Bereich von $\pm 30^\circ$, so fallen diese bei Versuch 2 und 3 vollständig weg. Derselbe Versuch wurde auch im Reflexionsarmen Raum durchgeführt, wobei sich ähnliche Relationen wie bei der im Studio durchgeführten ersten Versuchsreihe ergaben. Dabei war die Lokation insgesamt schlechter. Daraus kann man schließen, dass die Ortungsbewegungen, die Menschen unbewusst machen, sehr wichtig für die Lokalisation sind und offensichtlich auch gut simuliert werden können.

Abbildung 6: Messaufbau Mackensen



5 Diskussion und Ausblick

Die durchgeführte Arbeit zeigt den deutlichen Beitrag der HRTF auf die Lokalisationsfähigkeit des menschlichen Gehörs und bot die Gelegenheit, theoretische Kenntnisse in einer praktische Arbeit hörbar zu machen. Die im Selbstversuch Lokalisation war erstaunlich gut und wäre für einfache Anwendungen der virtuellen Quellpositionierung sicherlich verwendbar (Hörspiel etc.).

Die Messung der HRTF stellte sich jedoch besonders für die Versuchsperson als eher aufwändig und anstrengend heraus und ist für einen praktischen Gebrauch allein durch die benötigte Zeit sicher nicht geeignet. Hier wäre eine weitergehende Untersuchung schnellerer Methoden und deren Qualität interessant (z.B. reziproke Methode).

6 Literaturverzeichnis, Internetquellen

[1] SEEBER, B.: „*Untersuchung der Lokalisation in reflexionsarmer Umgebung und bei virtueller akustischer Richtungsdarbietung mit einer Laser-Pointer-Methode*“; Tagungsband Fortschritte der Akustik - DAGA 2002, Bochum, pp. 482–483, 2002
<http://www.mmk.ei.tum.de/publ/pdf/02/02see4.pdf>

[2] MINNAAR, PAULI, OLESEN, KRARUP, CHRISTENSEN, FLEMMING; MÖLLER, OLESEN: „*Localization with Binaural Recordings from Artificial and Human Heads*“; JAES Volume 49 Issue 5 pp. 323-336; May 2001

[3] UNBEKANNTER AUTOR: Hörspiel „*The Virtual Barbershop*“;
<http://ccgi.bluerabbit.plus.com/virtualbarbershop/>

[4] MAJDAK, BALAZS, LABAK: „*Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions*“, JAES., Volume 55, No. 7/8, pp. 623-637, 2007 July/August.

[5] GULDENSCHUH: „*Messung und Untersuchung der Einüsse der Head-Torso-Winkel auf die Aussenohrübertragungsfunktionen*“, Projektarbeit IEM, April 2008

[6] LINDAU, WEINZERL: „*Zur Wahrnehmbarkeit von Diskretisierungen in der dynamischen Binauralsynthese*“, Beitrag zur 25. Tonmeistertagung, Leipzig, 2008

[7] MACKENSEN: „*Head movements, an additional cue in localization*“, Dissertation TU Berlin, April 2004

7 Abbildungsverzeichnis

Abbildung 1: <http://games.cs.uni-magdeburg.de/audio/>

Abbildung 2: online.sfsu.edu/~psych200/unit6/cone.gif

Abbildung 3 und Daten von Zeichnungen 1 bis 6: <http://www.gras.dk/00012/00330/>

Abbildung 4: Projektarbeit Guldenschuh

Abbildung 5: Projektarbeit Lindau / Weinzerl

Abbildung 6: Diplomarbeit Mackensen

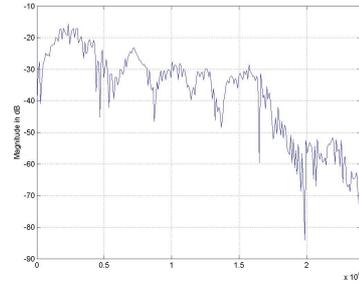
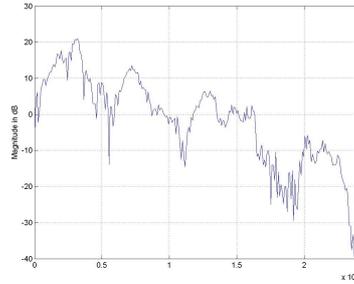
Anhang: Die Messergebnisse

Messpunkt

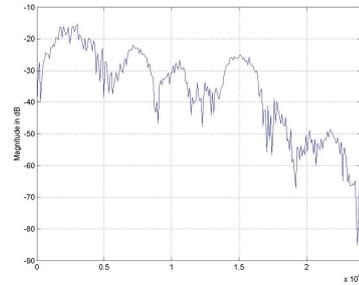
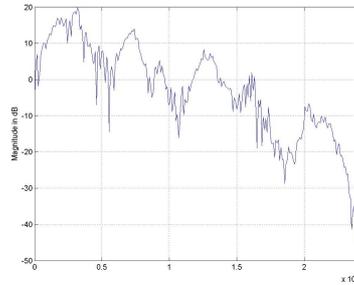
Linkes Ohr

Rechtes Ohr

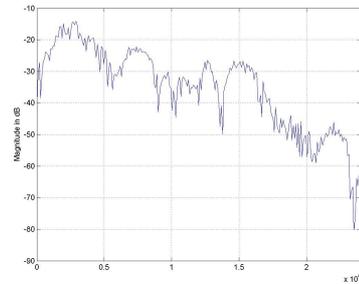
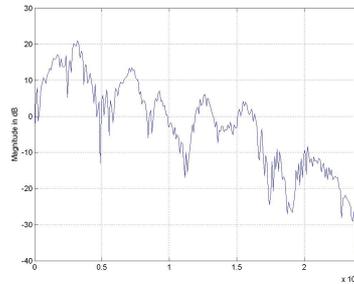
0 Grad



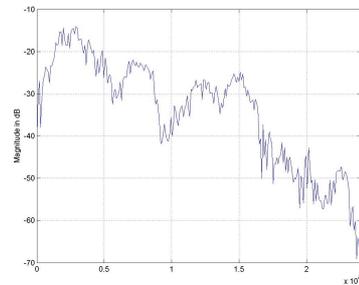
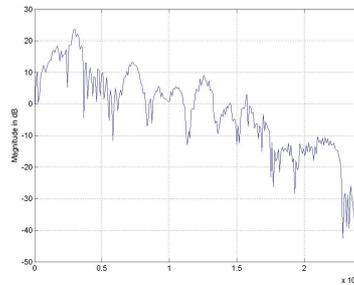
10 Grad



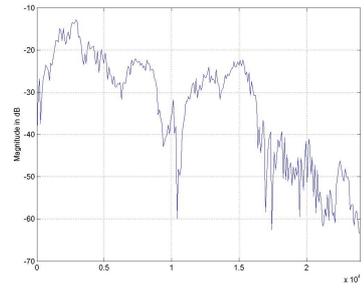
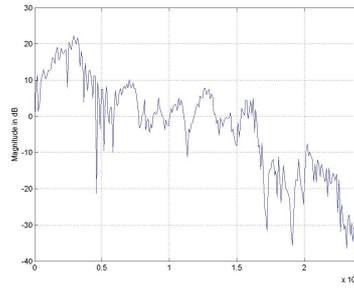
20 Grad



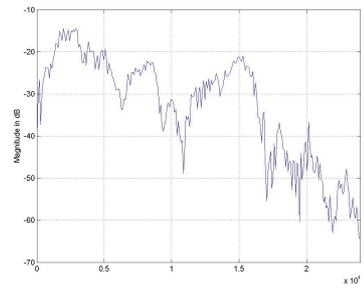
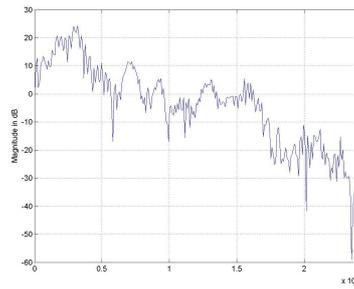
30 Grad



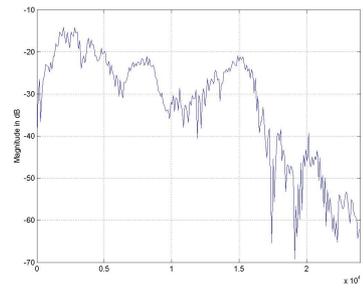
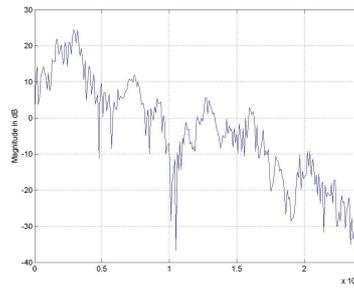
40 Grad



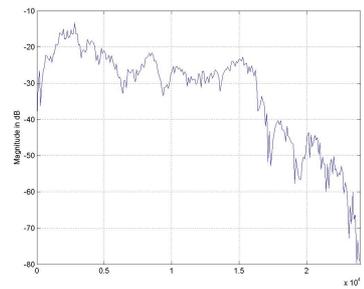
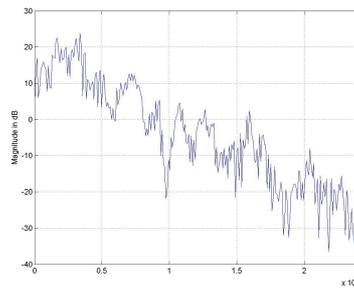
50 Grad



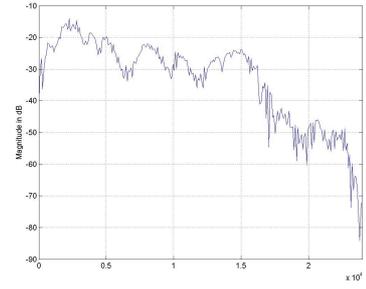
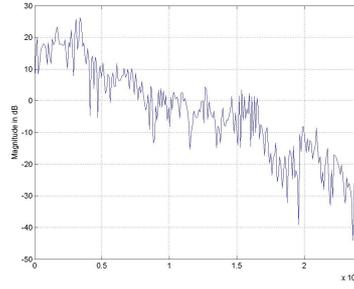
60 Grad



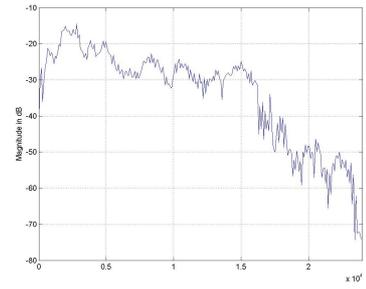
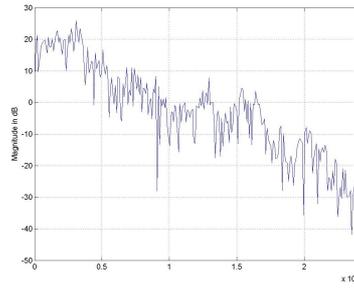
70 Grad



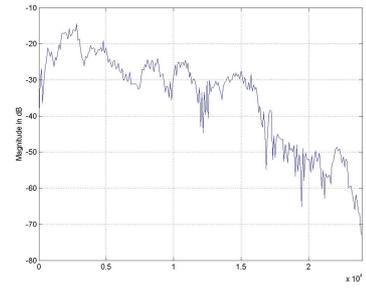
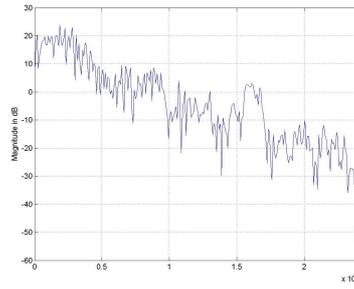
80 Grad



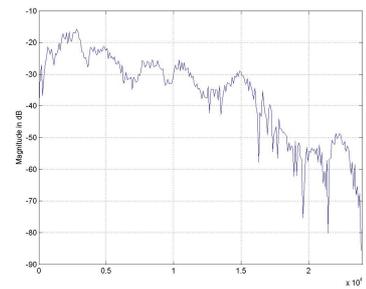
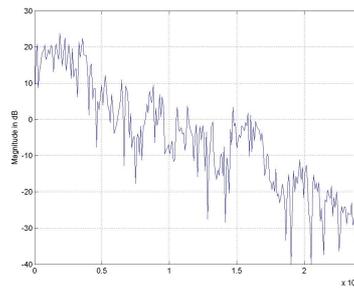
90 Grad



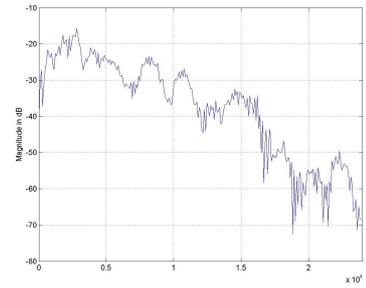
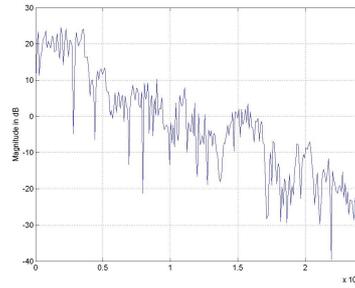
100 Grad



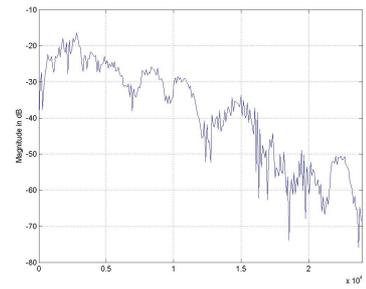
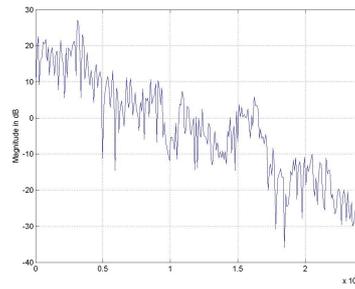
110 Grad



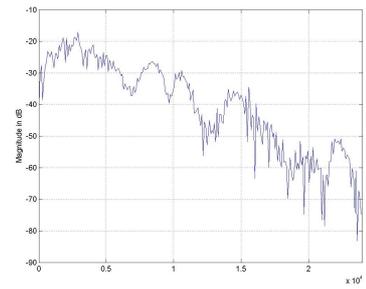
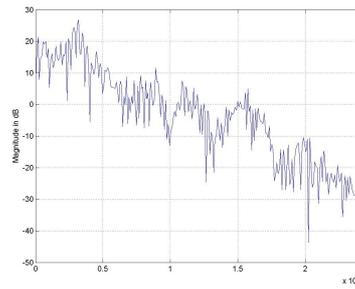
120 Grad



130 Grad



140 Grad



160 Grad

