

# Psychoakustische Messungen zur Diskrimination manipulierter Außenohrübertragungsfunktionen

Jörn Otten, Birger Kollmeier

AG Medizinische Physik, Carl-von-Ossietzky-Universität, D-26111 Oldenburg  
jotten@medi.physik.uni-oldenburg.de

## 1. Einleitung

HRTFs (Head Related Transfer Functions) beschreiben den grundlegenden Anteil der physikalischer Parameter, die der Mensch zur Lokalisation akustischer Quellen verwendet. Messungen von HRTFs zeigen deutlich, dass die Übertragungsfunktionen sehr individuell sind und somit Kunstköpfe, die einen mittleren Kopf nachbilden, nur bedingt geeignet sind individuell gemessene HRTFs zu ersetzen [1, 2, 3]. In diesem Zusammenhang läßt sich die Frage stellen, in wie weit HRTFs von individuell gemessenen HRTFs abweichen dürfen, ohne das die Abweichung einen perzeptiv relevanten Einfluss auf die Darbietung hat. In dieser Arbeit werden Experimente vorgestellt, die die Detektionsleistungen von Störungen an individuell gemessenen HRTFs beschreiben, um somit die perzeptiv zulässigen Abweichungen von der individuell gemessenen HRTF zu bestimmen. Als Störungen wurden gewählt: 1. Glättung des monauralen Spektrums. 2. Änderung der interauralen Pegeldifferenz.

## 2. Methode

In einem an Kulkarni et al. [4] angelehnten 2 Intervall, 2 AFC Verfahren, wurden den Versuchspersonen mittels Kopfhörer (AKG 501) Stimulussequenzen dargeboten. Diese Sequenzen bestanden aus 4 weißen Rauschen (bzw. Click-Trains), die in je zwei Intervallen zusammengefasst und durch eine Pause von einer Sekunde getrennt waren. In einem der beiden Intervalle wurde die HRTF, die die Sequenz in die jeweilige Richtung projiziert, eine der oben genannten Störungen zugefügt. Die Aufgabe der Versuchsperson war es, das Intervall zu identifizieren, in dem die HRTF des Stimulus manipuliert worden war. Die Sequenzen wurden aus der Horizontalebene für Azimute von  $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ$  dargeboten. Im folgenden werden die eingeführten Manipulationen kurz beschrieben.

### 1: Cepstral smoothing

Die Spektren der HRTFs weisen normalerweise eine starke Feinstruktur auf. Es stellt sich die Frage, in wie weit die Feinstruktur entfernt werden kann, ohne das diese Änderung einen Einfluß auf die Darbietung der virtuellen Schallquelle hat. Um diese Frage zu beantworten wurden die spektralen Feinheiten mittels einer Glättung entfernt. Dazu wurde das Spektrum der HRTF auf folgende Weise über eine Fourier-Reihe rekonstruiert (vgl. [4]):

$$\log(|H'(k)|) = \sum_{n=0}^M C(n) \cos(2 * \pi * n * k / N)$$

Das logarithmische Betragsspektrum von H wird über eine Fourierreihe angenähert, wobei sich die Koeffizienten  $C(n)$  aus dem Cepstrum der ursprünglichen HRTF ergeben.

Eine vollständige Rekonstruktion würde sich für  $M = N/2$  einstellen, wobei N die Anzahl an Filterkoeffizienten der unmodifizierten Impulsantwort angibt. Für  $M < N/2$  rekonstruiert die Fourier-Reihe das Spektrum nicht vollständig, da die Feinheiten der HRTF durch Cosinus Terme höherer Quefrenz repräsentiert werden.

### 2. ILD Morphing

Die zweite Manipulation bestand darin, die Pegeldifferenz der HRTF über den Austausch der Grobstruktur der individuellen HRTF zu ändern. Um spektral sinnvoll verteilte Pegeldifferenzen zu erhalten, wurde die individuelle Grobstruktur durch die spektrale Grobstruktur eines Kunstkopfes ersetzt. Das Betragsspektrum der variierten HRTF  $H'$  setzt sich damit nach der Formel zusammen:

$$|H'| = (1 - \alpha) |H| + \alpha \frac{|H|}{|H_s|} |H_{DH}|,$$

wobei H die ursprüngliche HRTF,  $H_s$  die Grobstruktur dieser HRTF und  $H_{DH}$  die Grobstruktur des Kunstkopfes beschreibt. Der Faktor  $\alpha$  ist ein Maß für die Mischung von individueller Struktur und der Grobstruktur des Kunstkopfes. Der Austausch wurde jeweils getrennt für das linke und rechte Ohr durchgeführt. Für  $\alpha=0$  ergibt sich keine Änderung, während für  $\alpha=1$  die Grobstruktur komplett ausgetauscht wird.

Um zu verhindern, dass die Versuchspersonen primär die Änderung der Klangfarben detektieren, wurden die vier weißen Rauschen einer jeden Stimulussequenz in 1/6 Oktavbändern mit zufälligen Verstärkungsfaktoren im Bereich von  $\pm 5$  dB spektral gefärbt. Dies gilt jedoch nur für die Änderung der Pegeldifferenzen und nicht für die spektrale Glättung.

## 3. Ergebnisse

In der Abbildung 1 sind die Detektionsleistungen über die effektiv auftretende Abweichung von der ILD aufgetragen. Jedes Symbol steht für eine Versuchsperson mit den Detektionsleistungen für Glättungen von 16, 32 und 64 Koeffizienten der Fourierreihe. In dem oberen Diagramm dieser Abbildung wurde weißes Rauschen als Stimulus verwendet. Die waagrecht eingezeichneten Linien geben das 5% Signifikanzniveau an. Es ist zu erkennen, daß schon mittlere Abweichungen von etwas mehr als 1dB auch bei  $90^\circ$  gut zu detektieren sind. Glättung mit 64 Koeffizienten sind für die Schalleinfallrichtung von  $90^\circ$  nur in Ausnahmefälle signifikant detektiert worden, während bei 32 Koeffizienten schon die Mehrzahl von Versuchspersonen signifikante Detektionsleistungen erzielt. In dem unteren Diagramm der Abbildung 1 sind die

### Spektrale Glättung

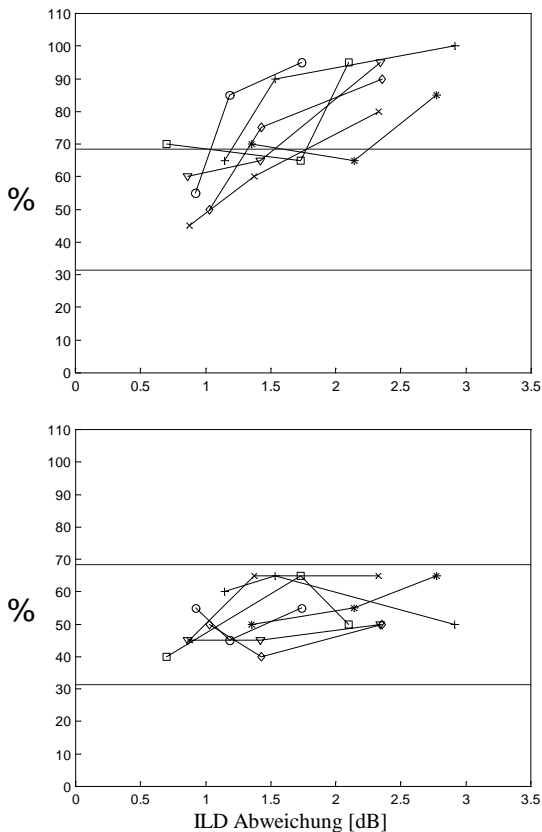


Abbildung 1:

Korrekt erkannte Störungen in Prozent als Funktion der effektiven Abweichung von der ILD bei Schalleinfall von 90°. Jedes Symbol steht für eine Versuchsperson mit 16, 32 und 64 Cepstralkoeffizienten. Oberes Diagramm: weißes Rauschen, unteres Diagramm: Click-Train.

Detektionsleistungen für den Click-Train als Stimulus aufgetragen. Man erkennt, daß es für 90° Schalleinfallrichtung nicht signifikant möglich war, die geglättete von der ungeglätteten HRTF zu unterscheiden. Dies ist ein Hinweis darauf, daß die Versuchspersonen die Änderung der Klangfarbe, die durch die Glättung eingeführt wurde, als Detektionscues verwendet haben. In der Abbildung 2 sind die Detektionsleistungen der Versuchspersonen über die durch den Austausch der Grobstrukturen entstandenen Pegeldifferenzen für Schalleinfall von 45° (oben) und 135° (unten) aufgetragen. Die fünf durch eine Linie verbundenen Symbole stehen jeweils für die Detektionsleistungen einer Versuchsperson mit  $\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$ . Durch die individuelle Struktur der HRTFs sind die durch die Manipulation entstandenen Abweichungen von der ursprünglichen ILD sehr unterschiedlich. Man kann aber auch hier erkennen, dass mittlere Abweichungen von etwas über 1 dB schon ausreichen, um detektierbar zu sein. Die Abhängigkeit der Detektionsleistung von der Abweichung der ILD variiert aber sowohl über die Versuchspersonen, als auch über die verschiedenen Schalleinfallswinkel sehr stark. In dem unterem Diagramm der Abbildung 2 (135° Schalleinfallrichtung), ist zu erkennen, dass alle Versuchspersonen schon Änderungen der ILD von unter 1 dB signifikant erkennen können. Da für diese Versuchsreihe die vier weißen Rauschen einer Stimulussequenz spektral unterschiedlich gefärbt waren, kann man davon ausgehen, dass die Detektionsleistungen anhand von Änderungen in der Lokalisation erzielt wurden.

### ILD Morphing

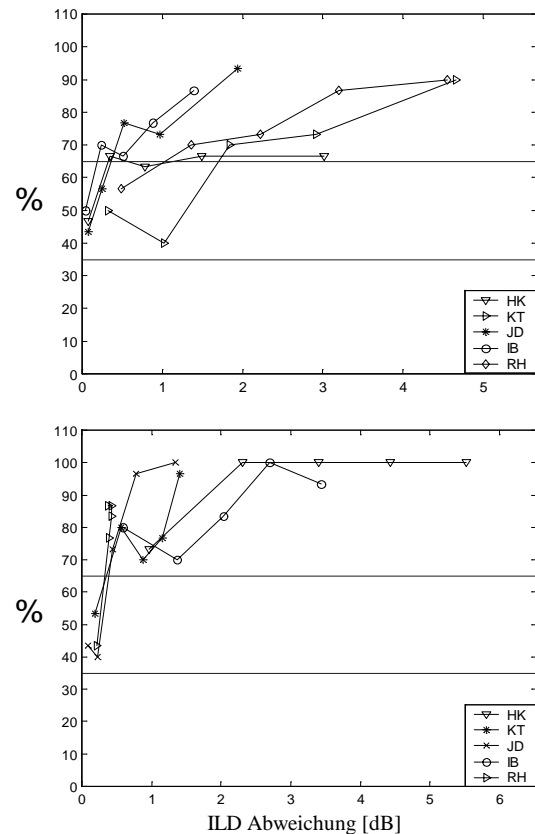


Abbildung 2:

Korrekt erkannte Störungen in Prozent als Funktion der effektiven Abweichung von der ILD. Jedes Symbol steht für eine Versuchsperson mit  $\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$ . Oberes Diagramm 45° Schalleinfallrichtung, unteres Diagramm 135° Schalleinfallrichtung

### 4. Fazit:

Es wurden Experimente zur Detektion von manipulierten HRTFs durchgeführt. Die Reduktion von Details in den Spektren der HRTFs hat für die hier vorgestellten Richtungen hauptsächlich zur Änderung der Klangfarbe beigetragen, die durch den Rauschträger auch sehr leicht zu detektieren ist. Der Click-Train verhindert größtenteils die Detektion der Klangfarbenänderung. Die Änderung der ILDs durch den Austausch der Grobstrukturen konnte anhand einer räumliche Veränderung detektiert werden. ILD-Abweichung von weniger als 1 dB sind signifikant detektiert worden, wobei die Detektionsleistung jedoch wesentlich von der Schalleinfallrichtung abhängt.

### 5. Literatur

1. Wenzel, E.M., Arruda, M., Kistler, D.J. (1993), „Localization using non-individualized head related transfer functions“, JASA 56
2. Wightman, F.I., Kistler, D.J. (1989), „Headphone simulation of freefield listening: I Stimulus synthesis“, JASA 85
3. Møller, H. Sorensen, M. F., Hammershoi, D., Jensen, C. B. (1995) „Head-Related Transfer Functions of Human Subjects“. J. Audio Eng. Soc., 43 (5)
4. Kulkarni, A., Colburn H.S., (1998), „Role of spectral detail in sound-source localization“. Nature 396