

**Algorithmen zur  
richtungsselektiven Verarbeitung  
von Schallsignalen -  
die Realisierung eines binauralen  
Cocktail-Party-Processor-Systems**

DISSERTATION  
zur Erlangung des Grades eines  
Doktor-Ingenieurs  
der Fakultät für Elektrotechnik  
an der Ruhr-Universität Bochum

von  
HARALD SLATKY  
Gelsenkirchen

Bochum 1992



## **Danksagung**

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für allgemeine Elektrotechnik und Akustik (AEA) an der Ruhr-Universität Bochum. Dem Inhaber des Lehrstuhl, Herrn Prof. Dr.-Ing. J. Blauert, der diese Arbeit ermöglichte, möchte ich an dieser Stelle herzlich danken.

Mein besonderer Dank gilt auch meinen ehemaligen Kollegen Dr.-Ing. Markus Bodden, Dr.-Ing. Werner Gaik, Priv.-Doz. Dr.-Ing. Herbert Hudde und Dr.-Ing. Siegbert Wolf sowie allen Mitarbeitern und Studenten am Lehrstuhl AEA, die mich durch ihre freundschaftliche Zusammenarbeit und konstruktive Kritik unterstützten.

Zu guter Letzt danke ich Herrn Prof. Dr.-Ing. U. Heute für sein freundliches Interesse an dieser Arbeit und für die Übernahme des Korreferats.



# Inhalt

1.	<b>Einleitung</b> .....	1
2.	<b>Psychoakustische Grundlagen</b>	
2.1.	Erkennen von Richtungen.....	5
2.2.	Verarbeiten von Signalen bestimmter Richtungen .....	7
2.3.	Merkmale und Leistungen der binauralen Signalverarbeitung .....	8
3.	<b>Hörversuche zur Lokalisation mehrerer Schallquellen</b>	
3.1.	Aufbau der Hörversuche.....	9
3.2.	Ergebnisse der Hörversuche .....	12
3.3.	Thesen zur Signalverarbeitung des Gehörs.....	22
3.4.	Vergleich Hörversuche - binaurale Modelle .....	23
4.	<b>Grundlagen der binauralen Signalverarbeitung</b>	
4.1.	Die Übertragungstrecke Schallquelle-Ohr.....	27
4.2.	Binaurale Informationen bei Anwesenheit mehrerer Schallquellen .....	31
4.3.	Algorithmen zur binauralen Verarbeitung der Ohrsignale .....	38
4.4.	Kreuzkorrelationsmodelle bei einer Schallquelle .....	39
4.5.	Die Kreuzkorrelationsfunktion bei mehreren Schallquellen.....	39
4.6.	Das komplexe Kreuzprodukt .....	44
5.	<b>Algorithmen zur Auswertung interauraler Phasenunterschiede ("Phasendifferenz-Cocktail-Party-Prozessor")</b>	
5.1.	Forderungen an einen Cocktail-Party-Prozessor.....	46
5.2.	Das interaurale Kreuzprodukt.....	46
5.3.	Der Phasendifferenz-Cocktail-Party-Prozessor .....	48
5.3.1.	Das interaurale Kreuzprodukt bei zwei Schallquellen .....	48
5.3.2.	Der Cocktail-Party-Prozessor-Algorithmus.....	49
5.3.3.	Signalverarbeitungs-Beispiele .....	51
5.3.4.	Eine Schallquelle mit veränderlicher Amplitude .....	55
5.4.	Komplexe Schallfelder.....	57
5.4.1.	Das interaurale Kreuzprodukt bei beliebigen Quellen .....	57
5.4.2.	Zwei Schallquellen mit zeitveränderlichen Amplituden.....	59
5.4.3.	Mehr als zwei Schallquellen mit konstanter Amplitude.....	61
5.4.4.	Diffuses Schallfeld.....	61
5.5.	Dominante Quellen.....	64
5.6.	Abbildung von Quellenschätzern auf eine gewünschte Richtung.....	65
5.7.	Zusammenfassung .....	71

<b>6.</b>	<b>Algorithmen zur Auswertung interauraler Pegelunterschiede ("Pegeldifferenz-Cocktail-Party-Prozessor")</b>	
6.1.	Problemstellung.....	74
6.2.	Bestimmung von Schallquellen-Parametern aus der Analyse interauraler Pegeldifferenzen.....	74
6.2.1.	Eine Schallquelle.....	74
6.2.2.	Zwei Schallquellen.....	75
6.2.3.	Mehr als zwei Schallquellen.....	76
6.2.4.	Diffuses Schallfeld.....	77
6.2.5.	Dominante Quellen.....	79
6.3.	Abbildung von Schätzern auf die gewünschte Richtung.....	81
6.4.	Verknüpfung von Schätzern aus interauralen Zeit- und Pegeldifferenzen.....	83
6.5.	Ausblick: Verbindung unterschiedlicher räumlicher Analyse- und Verarbeitungsverfahren.....	86
<b>7.</b>	<b>Ein Signalverarbeitungs-Rahmen für binaurale Modelle</b>	
7.1.	Verarbeitung der Eingangssignale.....	87
7.1.1.	Frequenzgruppen-Filter.....	87
7.1.2.	Erzeugung des analytischen Zeitsignals.....	89
7.1.3.	Datenreduktion.....	89
7.2.	Verarbeitung der Ausgangssignale.....	91
7.2.1.	Anforderungen an ein Resynthese-Verfahren.....	91
7.2.2.	Anpassung der Eingangssignale an Signalschätzer.....	92
7.2.3.	Verringern der Anzahl der Ausgangskanäle.....	95
7.2.4.	Bildung der Zeitfunktion aus Abtastwerten des analytischen Zeitsignals.....	95
7.3.	Gesamtdarstellung des Cocktail-Party-Prozessor-Modells.....	96
7.3.1.	Modellstruktur.....	96
7.3.2.	Leistungsfähigkeit der Cocktail-Party-Prozessoren.....	97
<b>8.</b>	<b>Steuerung des Cocktail-Party-Prozessors</b>	
8.1.	Detektionskriterien für Richtungsinformationen: Das Hören in Räumen.....	102
8.2.	Dynamische Effekte der Richtungserkennung: Der Präzedenz-Effekt.....	103
8.3.	Beschreibung des Präzedenz-Effekts durch ein binaurales Cocktail-Party-Prozessor-Modell.....	104
8.4.	Konsequenzen für die Steuerung eines Cocktail-Party-Prozessors.....	106
8.5.	Von der Prozessor-Steuerung zum Präzedenz-Prozessor.....	107
<b>9.</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>108</b>
	Anhang A: Auswerteverfahren für Hörversuche.....	112
	Anhang B: Frequenzgruppen-Modelle.....	113
	Anhang C: Vereinfachte Freifeld-Außenohr-Übertragungsfunktionen.....	117
	Anhang D: Ein flexibles Filterverfahren im Frequenzbereich.....	118
	Anhang E: Weitere Algorithmen zur Lösung des Cocktail-Party-Prozessor-Problems.....	123

Anhang F: Eine Programmstruktur für komplexe Prozesse (Parallelstrukturen auf sequentiellen Rechnern).....	125
Anhang G: Literatur .....	129

## Verwendete Symbole

### Indices

a	Signal a
abt	Abtastung, abgetastet
b	Signal b
c	Signal c
d	Diffusfeld
f	Freifeld-Übertragungsfunktion
FG	Frequenzgruppe
G	Gewichtungsfaktor
HE	Hörereignis
i	Summationsindex (z.B. Frequenzlinien)
k	Summationsindex (z.B. Frequenzlinien)
korr	Korrigierter Wert
l	linkes Ohr
m	auf den Bezugspunkt Kopfmitte bezogen
max	maximal vorkommender Wert
min	minimal vorkommender Wert
o	obere Grenze der Frequenzgruppe
p	Summationsindex für Schallquellen
q	Schallquelle (auch Summationsindex)
r	rechtes Ohr
soll	Wert der Sollrichtung
SQ	Schallquelle
u	untere Grenze der Frequenzgruppe
x	Parameter eines beliebigen Schätzers x (x= a oder b)
$\mu$	Mittelwert
$\sigma$	Standardabweichung
$\theta$	Einfallswinkel

### Exponenten

'	Schätzer
+	Kreuzleistungsdichte eines Signals
*	konjugiert komplex



## Formelzeichen

$a(t)$	Zeitfunktion des Schallsignals $a$
$a_m(t)$	Zeitfunktion des Schallsignals $a$ am Bezugspunkt Kopfmitte
$a_m'(t)$	Schätzer für den auf Kopfmitte bezogenen Betrag des Schallsignals
$\underline{a}_m(t)$	analytische Zeitfunktion des Schallsignals $a$ am Bezugspunkt Kopfmitte
$\underline{A}(f)$	Fourier-Transformierte des Schallsignals $a$
$\underline{A}_m(t)^2$	Quellenvektor (Bezugspunkt Kopfmitte)
$\underline{A}_m'(t)^2$	Quellenschätzer (Bezugspunkt Kopfmitte)
$\underline{A}_m(f, \tau)$	auf Kopfmitte bezogene Fourier-Transformierte des Schallsignals $a$
$\underline{A}_m^+(f, t)^2$	spektrale Kreuzleistungsdichte des Schallsignals $a(t)$ am Bezugspunkt Kopfmitte
AntwB	Antwortbereich für Hörversuche (z.B. $\pm 90^\circ$ )
$b(t)$	Zeitfunktion eines Schallsignals $b$
$b_m(t)$	Zeitfunktion des Schallsignals $b$ am Bezugspunkt Kopfmitte
$b_m'(t)$	Schätzer für den auf Kopfmitte bezogenen Betrag des Schallsignals
$\underline{b}_m(t)$	analytische Zeitfunktion des Schallsignals $b$ am Bezugspunkt Kopfmitte
$\underline{B}(f)$	Fourier-Transformierte des Schallsignals $b$
$\underline{B}_m(t)^2$	Quellenvektor (Bezugspunkt Kopfmitte)
$\underline{B}_m'(t)^2$	Quellenschätzer (Bezugspunkt Kopfmitte)
$\underline{B}_m(f, \tau)$	Fourier-Transformierte des Schallsignals am Bezugspunkt Kopfmitte
$\underline{B}_m^+(f, \tau)^2$	spektrale Kreuzleistungsdichte des Schallsignals $b(t)$ am Bezugspunkt Kopfmitte
$c_{\text{schall}}$	Schallgeschwindigkeit
$d$	Mikrofonabstand
$\underline{d}(t)$	interaurale Differenz
$E_m'^2$	Leistungsdichte im diffusen Schallfeld
$\underline{e}_m\theta$	analytisches Zeitsignal von Spiegelschallquellen
$f$	Frequenz
$f_{\text{abt}}$	Abtastfrequenz
$\underline{f}_{\text{korr}}$	Korrekturfaktor
$f_m$	Mittenfrequenz des Frequenzgruppen-Filters
$f_{\text{min}}$	untere Grenze des Übertragungsbereiches
$f_{\text{max}}$	obere Grenze des Übertragungsbereiches
$f_o$	obere Grenzfrequenz
$f_{oi}$	obere Grenzfrequenz der Frequenzgruppe $i$
$f_u$	untere Grenzfrequenz
$f_{ui}$	untere Grenzfrequenz der Frequenzgruppe $i$
$f()$	Funktion von
$F$	Fangbereich um eine Schallquelle, in dem Hörereignisse als korrekt lokalisiert gelten
$F$	Fourier-Transformierte

$F^{-1}$	inverse Fourier-Transformierte
FG	Frequenzgruppe
$g$	Bewertungsfaktor
$\underline{G}_m$ ' <sup>2</sup>	Ausgleichs-Signal für Korrekturverfahren
$h_l(t, \tau)$	Außenohr-Impulsantwort des linken Ohres
$h_r(t, \tau)$	Außenohr-Impulsantwort des rechten Ohres
$\underline{H}_l(f, \tau)$	Übertragungsfunktion linkes Ohr - "Kopfmitte"
$\underline{H}_{lf}(f, \tau)$	Freifeld-Außenohr-Übertragungsfunktion des linken Ohres
$\underline{H}_m(f, \tau)$	Freifeld-Übertragungsfunktion Kopfposition - "Kopfmitte"
$\underline{H}_{qk}(f, \tau)$	Freifeld-Übertragungsfunktion Quelle - Kopfposition
$\underline{H}_{ql}(f, \tau)$	Übertragungsfunktion Quelle - linkes Ohr
$\underline{H}_{qr}(f, \tau)$	Übertragungsfunktion Quelle - rechtes Ohr
$\underline{H}_r(f, \tau)$	Übertragungsfunktion rechtes Ohr - "Kopfmitte"
$\underline{H}_{rf}(f, \tau)$	Freifeld-Außenohr-Übertragungsfunktion des rechten Ohres
$\underline{H}_{rl}(f, \tau)$	interaurale Übertragungsfunktion
HE	Hörereignis
$\underline{k}(t)$	Kreuzprodukt
$l(t)$	reelle Zeitfunktion des linken Ohrsignals.
$\underline{l}(t)$	analytisches Zeitsignal des linken Ohrsignals.
$\underline{L}(f)$	Fourier-Transformierte des linken Ohrsignals.
LG	Lokalisationsgrad: Grad der Übereinstimmung von Schallereignis- und Hörereignis-Richtung
Lt	Lateralisation (-10=links, 0=mitte, 10=rechts)
$n_{oi}$	Flankensteilheit der hochfrequenten Filterflanke der Frequenzgruppe $i$
$n_{ui}$	Flankensteilheit der niederfrequenten Filterflanke der Frequenzgruppe $i$
$M, N$	Gesamtanzahl
$N_{HE}$	Anzahl der Hörereignisse bei einem Versuch
$N_{SQ}$	Anzahl der Schallquellen
$r(t)$	reelle Zeitfunktion des rechten Ohrsignals.
$\underline{r}(t)$	analytisches Zeitsignal des rechten Ohrsignals.
$\underline{R}(f)$	Fourier-Transformierte des rechten Ohrsignals.
$s$	Signal der Sollrichtung
$s_m$	Signal der Sollrichtung am Bezugspunkt Kopfmitte
$s_m$ ' <sup>2</sup>	Schätzer für die Leistung des Signal der Sollrichtung am Bezugspunkt Kopfmitte
$\underline{s}_m$ ' <sup>2</sup>	Schätzer für den Quellenvektor der Sollrichtung (Bezugspunkt Kopfmitte)
$S_{rl}(\tau)$	Kreuzkorrelationsfunktion
$S_{rl}(\tau, t)$	gleitende Kreuzkorrelationsfunktion
$\underline{S}_{rl}(f, t)$	Fourier-Transformierte der gleitenden Kreuzkorrelationsfunktion
SE	Schallereignis

$t$	Zeit
$T$	Zeitintervall
$T_{\mu}$	Integrationszeit zur Bestimmung statistischer Parameter
Vers	Versuch (Nummer)
VP	Versuchsperson (Nummer)
$w_r$	Rate-Wahrscheinlichkeit
$w(t)$	Fensterfunktion
$\underline{W}(f)$	Fourier-Transformierte der Fensterfunktion
$W_x$	Bewertungsfunktion für den Schätzer $x$
$x$	beliebiger Schätzer ( $x = a$ oder $b$ )
$x_m'^2$	Leistung eines beliebigen Schätzers am Bezugspunkt Kopfmittle
$\underline{X}_m'^2$	beliebiger Quellenschätzer (Bezugspunkt Kopfmittle)
$\alpha$	interaurale Dämpfung
$\alpha'$	Schätzer für die interaurale Dämpfung
$\beta$	interaurale Phase
$\beta'$	Schätzer für die interaurale Phase
$\beta_x'$	interaurale Phase des beliebigen Quellenschätzers $\underline{X}_m'^2$
$\Delta f$	(Mitten-)Frequenzdifferenz bei den Hörversuchen
$\Delta f_L$	(Mitten-)Frequenzdifferenz, ab der eine Schallquelle korrekt lokalisierbar wird.
$\Delta f_{L2}$	(Mitten-)Frequenzdifferenz, ab der beide Schallquellen gleichzeitig korrekt lokalisierbar werden.
$\Delta f_{FG}$	Bandbreite einer Frequenzgruppe
$\Delta L$	interaurale Pegeldifferenz
$\underline{\mu}$	komplexer Mittelwert der Kreuzkorrelationsfunktionen oder des Kreuzprodukts
$\mu_l$	Mittelwert des quadratischen Betrags des linken Ohrsignals
$\mu_r$	Mittelwert des quadratischen Betrags des rechten Ohrsignals
$\Phi$	Signalphase
$\Delta\Phi_{ab}$	Differenz der Signalphasen der Signale $a$ und $b$
$\Omega$	Momentankreisfrequenz
$\underline{\sigma}$	komplexe Standardabweichung der Kreuzkorrelationsfunktion oder des Kreuzprodukts
$\sigma_l$	Standardabweichung des quadratischen Betrags des linken Ohrsignals
$\sigma_r$	Standardabweichung des quadratischen Betrags des rechten Ohrsignals
$\tau$	interaurale Zeitdifferenz
$\tau'$	Verschiebungsparameter (Kreuzkorrelationsfunktion)
$\tau_L$	Normierte interaurale Zeitdifferenz ( $\pm 90^\circ$ entspricht $\pm 625\mu s$ )
$\tau_{max}$	maximale interaurale Laufzeit
$\tau_m$	mittlere interaurale Laufzeit im betrachteten Frequenzbereich.
$\tau_o$	mittlere Laufzeit Quelle-Empfänger
$\theta$	Einfallswinkel